

717	GCCATAAGAAATTGACNAGGATGGATCTCAGTGGGCAACGGCGTTGGATTCCCCANGCA	776
2282	TGATGCTGCTTTGCGAGGGCGCTGCGGCATCCCCAGATGCGAGGCTGCAGATGATTCAGTTGA	2341
777	TGATGCTGCTTTGCGAGGGCGCTGCGGCATCCCCAGATGCGAGGCTGCAGATGATTCAGTTGA	836
2342	GGAAAGTGC--ACCTGGAGTCCGGGGCTGTGCA--GGAGATGGCTT--CTGTGCTCGGCACC	2397
837	GGAAAGTGC--ACCTGGAGTCCGGGGCTGTGCA--GGAGATGGCTT--CTGTGCTCGGCACC	896
2398	AACCCACA 2405	
897	ANCCACA 904	
RESULT 2		
BO920369	AGENCOURT_9918853 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6383791	
LOCUS	5', mRNA sequence.	
DEFINITION	BO920369	
ACCESSION	BO920369.1 GI:22335067	
VERSION	EST.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 917)	
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/ .	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgapbs-remail.nih.gov Tissue Procurement: DCTD/DTP cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LICM2577 row: m column: 08 High quality sequence stop: 558.	
FEATURES		
source	1..917	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/clone="IMAGE:6383791"	
	/tissue_type="carcinoma, cell line"	
	/lab_host="DH10B (phage-resistant)"	
	/clone_lib="NIH_MGC_40"	
	/note="Organ: prostate; Vector: pOTB7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."	
ORIGIN		
Query Match	22.5%; Score 698; DB 13; Length 917;	
Best Local Similarity	98.5%; Pred. No. 8.1e-139;	
Matches	715; Conservative 0; Mismatches 10; Indels 1; Gaps 1;	
Qy	1504 GACGAGGGGAGGGCGGGCGAGCCAGACACGAGACCTGACCGAGCTGTTCACCGAGTAC	1563
Db	1 GACGAGGGGAGGGCGGGCGAGCCAGACACGAGACCTGACCGAGCTGTTCACCGAGTAC	60
Qy	1564 GCGTTTTCTGAAAGAGCTTCTGGCACTACACAGCGCTTCTGTTGGACTCCTCAAC	
Db	61 GCGTTTTCTGAAAGAGCTTCTGGCACTACACAGCGCTTCTGTTGGACTCCTCAAC	120
Qy	1624 GAGGAGACCCAGGAGCCACTCGGAAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAG	1683
Db	121 GAGGAGACCCAGGAGCCACTCGGAAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAG	180
Qy	1684 ATGACCTGTCAGTGTATCCAAAGCAAGCTCAGAGCGAGCTCCACCTGCACGAG	1743
Db	181 ATGACCTGTCAGTGTATCCAAAGCAAGCTCAGAGCGAGCTCCACCTGCACGAG	240
Qy	1744 GGCCTCTGGAGTCTTCAGCTGCTGTACAGATCCAGGAGGAGGTTTATCCAGCAG	1803
Db	241 GGCCTCTGGAGTCTTCAGCTGCTGTACAGATCCAGGAGGAGGTTTATCCAGCAG	300
Qy	1804 GCCCTGAGCCACTTCCAGGTGATCGTGTTCAGCAACATTTGCTTCCAAAGATGGAGCACATG	1863
Db	301 GCCCTGAGCCACTTCCAGGTGATCGTGTTCAGCAACATTTGCTTCCAAAGATGGAGCACATG	360
Qy	1864 GTCTCTCTGTTCTGTCTGAAGCGTGTGAGGAGCGCCAGTGTGCTGCTATGCGGCC	1923
Db	361 GTCTCTCTGTTCTGTCTGAAGCGTGTGAGGAGCGCCAGTGTGCTGCTATGCGGCC	420
Qy	1924 ACCTACAGCGGGAGCGGAGACCGCGCAGGTGCTCCGAGGAGCGCACACGCTGTG	1983
Db	421 ACCTACAGCGGGAGCGGAGACCGCGCAGGTGCTCCGAGGAGCGCACACGCTGTG	480
Qy	1984 GTGAGCTTCACACAGAGAGACCGTCTGCTGGAAGCTTACAGTGAACATCTGGCAGCG	2043
Db	481 GTGAGCTTCACACAGAGAGACCGTCTGCTGGAAGCTTACAGTGAACATCTGGCAGCG	540
Qy	2044 GCCCTGTGCAACAATCCAAACCTGATAGAGTGTCTGTGTACCGAAATGCGCTGGCGAGC	2103
Db	541 GCCCTGTGCAACAATCCAAACCTGATAGAGTGTCTGTGTACCGAAATGCGCTGGCGAGC	600
Qy	2104 CGGGGGTGAAGCTGCTGTTCAGGACTCAGACACCCCACTGCAACTTCAGAACTTG	2163
Db	601 CGGGGGTGAAGCTGCTGTTCAGGACTCAGACACCCCACTGCAACTTCAGAACTTG	657
Qy	2164 AGGCTGAAGAGTGCGGATCTTCCAGCTCAGCTCCGAGGACCTCTCTGAGCTTCTATA	2223
Db	658 AGGCTGAAGA--GTGCGGATCTGAGCTCAGCTTCGAGGACCTCTCTGAGCTTCTATA	716
Qy	2224 GCCAATAAGAAATTTGACAAGGATGGATCTCAGT--GGCAACGGCGTGGATT--CCAGGCA	2281

[illegible]

```

RESULT 3
BQ900330
LOCUS
DEFINITION
AGENCOURT_8678096 NIH_MGC_40 Homo sapiens cDNA clone IMAGE:6381082
5', mRNA sequence.
BQ900330.1 GI:22292344
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 937)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLC#2570 row: 1 column: 11
High quality sequence start: 5
High quality sequence stop: 502.

```

```

AW673661/c
LOCUS          571 bp      mRNA      linear      EST J1-APR-2000
DEFINITION     ba56e01.y3 NIH MGC_10 Homo sapiens cDNA clone IMAGE:2900528.5
                similar to TR:075434 075434 ANGIOTENSIN/VASOPRESSIN RECEPTOR
                AII/AVP. ;, mRNA sequence.
ACCESSION      AW673661
VERSION        AW673661.1 GI:7538896
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
                1 (bases 1 to 571)
REFERENCE      NIH-MGC http://mgi.nci.nih.gov/
AUTHORS        National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE          Unpublished (1999)
JOURNAL
COMMENT        Other ESTs: ba56e01.x1
                Contact: Robert Strausberg, Ph.D.
                Email: cgabbs-x@mail.nih.gov
                Tissue Procurement: ATCC
                cDNA Library Preparation: Life Technologies, Inc.
                cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
                DNA Sequencing by: Washington University Genome Sequencing Center
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                image.llnl.gov/image/html/iresources.shtml

Possible reversed clone: similarity on wrong strand
Seq primer: -40RP from Gibco
High quality sequence stop: 426.
FEATURES             source
    1..571
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:2900568"
    /cell_line="MGC36"
    /lab_host="DH10B"
    /clone_lib="NIH MGC_10"
    /note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
    Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
    Average insert size 1.5 kb. Library prepared by Life
    Technologies."
ORIGIN
Query Match          17.4%; Score 539.8; DB 10; Length 571;
Best Local Similarity 97.7%; Pred. No. 5.6e-105;
Matches 558; Conservative 0; Mismatches 12; Indels 1; Gaps 1;

QY 1297 GGGTTGTCTCTTGGCGGAGATGGGCTCTGGAATC-AGAAATCCTATTGAGGAGCA 1355
DB      |||
QY 1356 GGACCTCCGGAAGCAGCGGCTAGACGGGGAAGACGCTCTGTGCTTCTCAACATGAACAT 1415
DB      |||
QY 511  GGACCTCCGGAAGCAGCGGCTAGACGGGGAACACGCTCTGTGCTTCTCAACATGAACAT 452
QY 1416 CTTCCAGGAAGGACATCAACTGTGAGAGTACTACAGCTTCTACCTTGTGAGTTCCAGGA 1475
DB      |||
QY 451  CTTCCAGGAAGGACATCAACTGTGAGAGTACTACAGCTTCTACCTTGTGAGTTCCAGGA 392
QY 1476 ATTCCTTTGAGCTATGTACTATATCTCTGACGAGGGAGGGCGGGGAGGGCCGAGCA 1535
DB      |||
QY 391  ATTCCTTTGAGCTATGTACTATATCTCTGACGAGGGAGGGCGGGGAGGGCCGAGCA 332
QY 1536 GGACGTGACCGGCTGTGACCGAGTACCGGTTTCTGAAAGAGAGTCTTCTGGCACTAC 1595
DB      |||
QY 331  GGACGTGACCGGCTGTGACCGAGTACCGGTTTCTGAAAGAGAGTCTTCTGGCACTAC 272
QY 1596 GAGCGCTTCCGTTTGGACTCTCTGACGAGGAGCCAGGAGCCACTCGAGAGAGTCT 1655
DB      |||
QY 271  CAGCGCTTCCGTTTGGACTCTCTGACGAGGAGCCAGGAGCCACTCGAGAGAGTCT 212

```

RESULT 5
 AG067278
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AG067278 664 bp DNA linear GSS 03-NOV-2001
 Pan troglodytes DNA, clone: PTB-057D11.R, genomic survey sequence.
 AG067278
 AG067278.1 GI:16619080
 GSS.
 Pan troglodytes (chimpanzee)
 Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library PTB
 Unpublished
 2 (bases 1 to 664)
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-chou, Tsukuba, Ibaraki, Japan, 305-8565, Japan
 (E-mail: chimpes@gs.riken.go.jp, URL: http://hgp.gs.riken.go.jp/
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library PTB. This BAC end
 was generated during the R&D process and may have higher chance of
 clone tracking errors.

PRIMERS
 Sequencing: M13Rev
 LIBRARY
 Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI.
 Location/Qualifiers
 1..664
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="PTB-057D11.R"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
 Query Match 16.7%; Score 518.8; DB 29; Length 664;
 Best Local Similarity 92.7%; Pred. No. 2e-100;
 Matches 555; Conservative 0; Mismatches 43; Indels 1; Gaps 1;
 QY 750 GCCTCTCAGGAGCTCATCCGAGTCCCGAGCGCTCTCTTTTCATCATCGAGGCTTCA 809
 DB |||
 QY 49 GCATGCTGAGAGCTCATCCGAGTTCGAGGCGCTCTTTTCATCATCGAGGCTTCA 108
 DB |||
 QY 810 TGAGCTCAAGCCCTCTTTTCCAGATCTCTCAGGAGCCCTGGTCCCTCTGTGGAGAGAA 869
 DB |||
 QY 109 TGAGCTCAAGCCCTCTTTTCCAGATCTCTCAGGAGCCCTGGTCCCTCTGTGGAGAGAA 168
 DB |||

QY 870 ACGGCCACGGAGCTGCTTTTAAACAGCTTAAATTCGGAAGAGCTGCTCCCTCAGCTATC 929
 Db |||||
 QY 169 ACGGCCACGGAGCTGCTTTTAAACAGCTTAAATTCGGAAGAGCTGCTTTTTCAGCTATC 228
 Db |||||
 QY 930 TTTGCTCATCACACAGGCCACCGGCTTTGGAGAGCTTCCACGCTGCTGTCGAGCACCC 989
 Db |||||
 QY 229 TTTGCTCATCACACAGGCCACCGGCTTTGGAGAGCTTCCACGCTGCTGTCGAGCACCC 288
 Db |||||
 QY 990 CAGGCATGTGGAGATCCTGGGCTTCTTGAGGAGAGCTTCAATTTAGTGAGGAGCAACGAGCTCT 1049
 Db |||||
 QY 289 CAGGCATGTGGAGATCCTGGGCTTCTTGAGGAGAGCTTCAATTTAGTGAGGAGCAACGAGCTCT 348
 Db |||||
 QY 1050 TTTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTTAGTGAGGAGCAACGAGCTCT 1109
 Db |||||
 QY 349 TTTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTTAGTGAGGAGCAACGAGCTCT 408
 Db |||||
 QY 1110 CTTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTTAGTGAGGAGCAACGAGCTCT 1169
 Db |||||
 QY 409 CTTTCCAAATGACAGAGCAGCGGCCCAAGTCTTCAATTTAGTGAGGAGCAACGAGCTCT 468
 Db |||||
 QY 1170 GCTGAGGCTGGGGGCTGTTGAGAGCAGAGCTCCAGGAGCACCACCTGAGCTGATGCT 1229
 Db |||||
 QY 469 GCTGAGGCTGGGGGCTGTTGAGAGCAGAGCTCCAGGAGCACCACCTGAGCTGATGCT 528
 Db |||||
 QY 1230 CTACTGCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1288
 Db |||||
 QY 529 CTACTGCTGAGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
 Db |||||
 QY 1289 ACCAGAGAGGCTGCTCTTGGGGGAGAGTGGGCTTGGAGTCAAGAAATCTTATT 1347
 Db |||||
 QY 589 TGCANAGAGGCTGCTCTTGGGGGAGAGTGGGCTTGGAGTCAAGAAATCTTATT 647
 Db |||||

RESULT 6

AW956628 568 bp mRNA linear EST 01-JUN-2000
 LOCUS EST168698 MAGG resequences, MAGD Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION AW956628
 VERSION AW956628.1 GI:8146311
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharap,S., Gaspard,R., Gay,C.,
 Holt,I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and
 Quackenbush,J.
 TITLE Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 JOURNAL Unpublished (2000)
 COMMENT The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 94
 Seq primer: Reverse.
 Location/Qualifiers
 1..568
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGD"
 /note="Vector: pBluescriptSKm"

FEATURES

source

Query Match 15.9%; Score 494.2; DB 10; Length 568;
 Best Local Similarity 97.4%; Pred. No. 3.5e-95;
 Matches 523; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

ORIGIN

1049764
 Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>)
 or from Research Genetics (<http://www.htcsc.washington.edu>)
<http://www.htcsc.washington.edu>
 Plate: 984 row: E column: 8
 Seq primer: SP6

QY 1638 CCACCTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGACATCAAGATGACCTGTGCA 1597
 Db |||||
 QY 1698 GTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTCCACCTGTCAGCAGGGCTCTCTTGGAGTT 1757
 Db |||||
 QY 61 GTGGATCCAAAGCAAAAGCTCAGAGCGACGGCTCCACCTGTCAGCAGGGCTCTCTTGGAGTT 120
 Db |||||
 QY 1758 CTTTACGCTCTTGTACAGATCCAGAGGAGGAGTTTATCCAGCAGGCCCTGAGCCACTT 1817
 Db |||||
 QY 121 CTTTACGCTCTTGTACAGATCCAGAGGAGGAGTTTATCCAGCAGGCCCTGAGCCACTT 180
 Db |||||
 QY 1818 CCAGGTGATCTGCTGTCAGCAACATTCCTCCAGATGGAGCAGATGCTCTCTCTCTCTG 1877
 Db |||||
 QY 181 CAGGTGATCTGCTGTCAGCAACATTCCTCCAGATGGAGCAGATGCTCTCTCTCTCTG 240
 Db |||||
 QY 1878 TCTGAAGCGCTGTCAGGAGCGGCCAGGTGCTGCACTTTGTATGGCGCCACCTACAGCGCGA 1937
 Db |||||
 QY 241 TCTGAAGCGCTGTCAGGAGCGGCCAGGTGCTGCACTTTGTATGGCGCCACCTACAGCGCGA 300
 Db |||||
 QY 1938 CGGGAGAGACCGCGCGAGGTGCTCCGAGAGGCGACACGCTGTTGGTGAGCTCAGACC 1997
 Db |||||
 QY 301 CGGGAGAGACCGCGCGAGGTGCTCCGAGAGGCGACACGCTGTTGGTGAGCTCAGACC 357
 Db |||||
 QY 1998 AGAGAGAGACCGCTGCTGTCAGCGCTTACAGTGAACATCTGGCAGCGCGCTGTGACACCA 2057
 Db |||||
 QY 358 AGAGAGAGACCGCTGCTGTCAGCGCTTACAGTGAACATCTGGCAGCGCGCTGTGACACCA 417
 Db |||||
 QY 2058 TCCAAACCTGATAGAGCTGCTCTGTACCGAAATGCCCTGGGCGAGCGGGGGTGAAGCT 2117
 Db |||||
 QY 418 TCCAAACCTGATAGAGCTGCTCTGTACCGAAATGCCCTGGGCGAGCGGGGGTGAAGCT 477
 Db |||||
 QY 2118 GCTCTGTCAGGAGCTCAGACACCGCCAACTGCAAACTTCAGAACTTCAGAA-CTGAGGCTGAGAG 2174
 Db |||||
 QY 478 GCTCTGTCAGGAGCTCAGACACCGCCAACTGCAAACTTCAGAA-CTGAGGCTGAGAG 533
 Db |||||

RESULT 7

AW9726243 496 bp DNA linear GSS 14-JUL-1999
 LOCUS HS_5408 A2 C04 SP6E RPCI-11 Human Male BAC Library Homo sapiens
 DEFINITION genomic clone SP6E-984 Col=8 Row=E, genomic survey sequence.
 ACCESSION AW9726243
 VERSION AW9726243.1 GI:5485912
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 496)
 AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
 Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
 Hood,L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering.bac.htm>)
 or from Research Genetics (<http://www.htcsc.washington.edu>)
<http://www.htcsc.washington.edu>
 Plate: 984 row: E column: 8
 Seq primer: SP6

QY 1960 TCCGAGGAGCGCAC 1976
 Db 481 TCTCGATGCGCAC 497

RESULT 9
 BI911853
 LOCUS
 DEFINITION 764 bp mRNA linear EST 16-OCT-2001
 mRNA sequence.
 ACCESSION BI911853
 VERSION BI911853.1 GI:16175736
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 764)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapb-rc@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM11534 row: c column: 10
 High quality sequence start: 22
 High quality sequence stop: 764.
 Location/Qualifiers
 1..764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5212737"
 /tissue_type="leukocyte"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_118"
 /notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV
 (destroyed); RNA source leukocytes from anonymous pool of
 non-activated adult donors. Library is oligo-dT primed
 and directionally cloned (EcoRV site is destroyed upon
 cloning). Average insert size 1.7 kb, insert size range
 1.2-3.3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 027. Note:
 this is a NIH_MGC Library."

ORIGIN
 Query Match 13.7%; Score 426; DB 12; Length 764;
 Best Local Similarity 86.7%; Pred. No. 1.9e-80;
 Matches 539; Conservative 0; Mismatches 0; Indels 83; Gaps 3;

QY 1 ATGCTACGACCGCGAGCGGAGCGGCTGTGCGCTGTCCACTTCTGGAAGACTC 60
 Db 143 ATGCTACGACCGCGAGCGGAGCGGCTGTGCGCTGTCCACTTCTGGAAGACTC 202

QY 61 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCGAGCGGCTGGGAGAA 120
 Db 203 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCGAGCGGCTGGGAGAA 262

QY 121 GCGAAGATCCCTGGGAGAGATGGAGAGCGGCTCCCTGTGAATGCCCAGCTGCTC 180
 Db 263 GCGAAGATCCCTGGGAGAGATGGAGAGCGGCTCCCTGTGAATGCCCAGCTGCTC 322

QY 181 ATCACCACCTCGGCGCAGAGGCGCTGAGGTTGGCTCTCAGACCTTTGAGCGGATA 240
 Db 323 ATCACCACCTCGGCGCAGAGGCGCTGAGGTTGGCTCTCAGACCTTTGAGCGGATA 382

QY 241 AACAGGAGGACCTGTGGGAGAGGACAGAGAGGACCTGTGTAGG----- 288
 Db 383 AACAGGAGGACCTGTGTGGGAGAGGACAGAGAGGACCTGTGTAGGATACCCACCT 442

QY 289 ----- 288

Db 443 GGTGGCCGCTCTCACTTGGGAACCACTCAACATGCTTCTGGAGTCTCTCTTGTGACT 502

QY 289 -----CATCCCAAGAAACCTACAGGACTATGTCCGCAAGAAATTCGGCTCATG 339
 Db 503 CCAAGAAAGATATCCCAAGAAACCTACAGGACTATGTCCGCAAGAAATTCGGCTCATG 562

QY 340 GAAGACCGCAATGCGG-CCTAGGGGAATGTCTCACTCAGCCAGCGGTACACCCGCT 398
 Db 563 GAAGACCGCAATGCGGCCCCCTAGGGGAATGTCTCACTCAGCCAGCGGTACACCCGCT 622

QY 399 CTTGCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGGCTAGCCCATCAAGATAGAGACCTCT 517
 Db 623 CTTGCTGTGAAGGAGCACTCAAAACCCCATGCAAGTCCAGGCTAGCCCATCAAGATAGAGACCTCT 742

QY 518 TTGAGCCAGAGGAGGCGGCC 539
 Db 743 TTGAGCCAGAGGAGGCGGCC 764

RESULT 10
 AQ443134
 LOCUS
 DEFINITION 449 bp DNA linear GSS 31-MAR-1999
 HS-5130_AI_B05_SP6E_RPCI-11 Human Male BAC Library Homo sapiens
 Genomic clone Plate=706 Col=9 Row=C, genomic survey sequence.
 ACCESSION AQ443134
 VERSION AQ443134.1 GI:4554473
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 449)
 AUTHORS Mahairas, G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,
 Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and
 Hood, L.
 TITLE Sequence-tagged connectors: A sequence approach to mapping and
 scanning the human genome
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
 MEDLINE 99380589
 PUBMED 10449764
 COMMENT Contact: Mahairas GG, Wallace JC, Hood L
 High Throughput Sequencing Center
 University of Washington
 401 Queen Anne Avenue North, Seattle, WA 98109, USA
 Tel: (206) 616-3618
 Fax: (206) 616-3887
 Email: jwallace@u.washington.edu
 Clones are derived from the human BAC library RPCI-11. For BAC
 library availability, please contact Pieter de Jong
 (pieter@dejong.med.buffalo.edu). Clones may be purchased from
 BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
 or from Research Genetics (<http://www.resgen.com>). BAC end Web Server:
<http://www.htsc.washington.edu>
 Plate: 706 row: C column: 9
 Seq primer: SP6
 Class: BAC ends
 High quality sequence stop: 449.
 Location/Qualifiers
 1..449
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"

FEATURES
 source

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahy, J., Helton, E., Ketterman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalls, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 3359)
 Strausberg, R.
 Direct Submission
 Submitted (14-JAN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Aug 19, 2003 this sequence version replaced gi:19204229.
 Contact: MGC help desk
 Email: cgapbs-femail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgion, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 39 Row: 9 Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 19745161
 This clone has the following problem: frame shifted.
 FEATURES
 Location/Qualifiers
 1..3359
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4811303"
 /tissue_type="Placenta, choriocarcinoma"
 /clone_lib="NIH MGC 21"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"

ORIGIN

Query Match 12.1%; Score 377; DB 11; Length 3359;
 Best Local Similarity 48.9%; Pred. No. 1.7e-69;
 Matches 1269; Conservative 0; Mismatches 1260; Indels 66; Gaps 7;
 QY 513 CCTCTTTGAGCAGACGAGGAGGCGCCCGAGCCACCGCGACCGTGGTCTATCGAGGCGC 572
 Db 469 CGCGCTTTTGTCTCCCAAGAAATCTGGGAAACAGCCACGTACAGTGAATTAATCAAGGACC 528
 QY 573 GGCAGGGATAGGCAAGTCCATGCTGGSCACACAGGTGATGCTGGACTGGCGGACGGGAA 632
 Db 529 ACAAGGAATTTGAAACACGACACTCTGTATGAGCTGATGATGGCTGTGTCGGACACAA 588
 QY 633 GCTCTTCCAGGAGATTGATATCTCTTCTTACATCAACTGCGGAGATGAACCCAGAG 692

589 GAICTTTTCGGATAGGTTCTCTGTACAGTTCTATTCTTCTGTCAGAGAACTGAGGGAG-- 646
 QY 693 TGGCAGCGAATGCGATGCAAGACCTCATCTTCAGCTGTGTGGCTGAGCCAGCGGCC 752
 Db 647 -TTGCCGCCAACGAGTTTGGCTGACTTGAATTTTCAGAGAGTGGCTGACCCCGTGTCTC 705
 QY 753 TCTCCAGAGCTCATCGAGTTCCGAGCGCTCTTTTCATCATCAACGGCTTCGATGA 812
 Db 706 TATAACAGAGATCGTGTCTCAACCGGAGAGACTCTTGTTCGTATCAGACAGCTTCGAAGA 765
 QY 813 GCTCAAGACCTTTCTTCCACGATCTCTCAGGACCTCTGCTGGGAGGAGAAACG 872
 Db 766 GCTGCAGGCGGCTTGAACGAAACCGGATCTGTGTGGTACTTGTATGAGAAACG 825
 QY 873 GCCACGGAGCTCTTCTTAAACAGCTTAATTCGGAAGAGCTCTCCCTGAGCTATCTTT 932
 Db 826 GCCGGTGCAGTGTCTTCTGAGCAGTTTGTCTGAGGAAGAGATGCTCCGAGGCTCCCT 885
 QY 933 GCTCATCACACAGCGGCCACGCTTTGGAGAGCTCCACCGTCTGCTGAGGACCCCGAG 992
 Db 886 GCTATCGCCATCAACCCGCTGTGCCGAAAGGAGCTCCGGGATCAGGTGACGATCTCAGA 945
 QY 993 GCATGTGGAGATCTTGGGCTTCTCTGAGGCAGAAAGGAAGAAATACTTCTACAAGTATTT 1052
 Db 946 AATCTACGACCCCGGGGATTCAACGAGAGTATAGTTAGTGTATTTCTGTCTTCTT 1005
 QY 1053 CCACATGACAGACGCGGCGCAAGTCTTCAATTAAGTGGAGGACAAAGAGCTCTCTT 1112
 Db 1006 CAAAGACCCGAAAGAGCCATGGAAGCTTCAATCTTTGAAGAGAAAGTGAACAGCTGTT 1065
 QY 1113 CACCATGTCTTCCCTCCCTGCTGGTGTGCTGGTGTGTTGTTACCTGCTCCAGCAGAGCT 1172
 Db 1066 TTCCATATGCCAATCCCGTCTCTGCTGGATCTCTGTGTACCACTGTAAGCAAGAGAT 1125
 QY 1173 GGAGGTGGGGGCTCTTGAGACAGACGCTCCAGGACCAACCACTGCACTGTACATGCTCTA 1232
 Db 1126 GCAGAAAGGAAAGACCTTGGCCCTGACCTGCCAGAGCACTACCTCTGTGTACTCTCTT 1185
 QY 1233 CTTGCTGAGTCTGATGCAACCAAG---CGGGGGGCGCCGCTCCAGCCCCCAACCA 1289
 Db 1186 CGTCTTTAACTGTTTCAACCTGAGGCTGCGAGGCGCCGACCTCCGAAACCCAGCACCA 1245
 QY 1290 CAGAGAGGCTGTGCTCTCTGCGGACAGATGGCTCTCGAATCAGAAATCTATTGTA 1349
 Db 1246 GCTGAAGGCCCTGTGCTCTCTGCTCAGAGGATGTTGGACAGACACATTTGAGTTTG 1305
 QY 1350 GGAGCAGGACCTCCGGAAGCAGCGCTTAGACGGGGAAGACGCTCTGTGCTTCTCAACAT 1409
 Db 1306 TGAAGACGACCTCCGAGAAATGGGGTTGTTGACGCTGACATCCCTGCGCTGCTGGGCAC 1365
 QY 1410 GAACATCTTCCAGAGGACATCACTGTGAGAGTACTACAGCTTCACTCCCTGAGTTT 1469
 Db 1366 CAAGATACTCTGAAGTACGGGAGCGTGAAGAGCTCTACGTTCTCTCCACGCTGTAT 1425
 QY 1470 CAGAGATTTCTTGGAGCTATGATATATCTTGGACGAGGGGGGAGGGGCGGCGCC 1529
 Db 1426 CCAGGAGTTCTGTGCGGCTTGTCTATTTGTTCAAGAGCCACCTTGTATCTCTCACCC 1485
 QY 1530 AGACCAG-----GACGTGACCCAGGCTGTTGACCGAGTACGGCTTTCTGAAAGGAGCTT 1583
 Db 1486 AGCTGTGAGATGTGTACAGGAATTCGTAGTTCCTCAATTTTGAAGAGCAAGAGAGACCA 1545
 QY 1584 CTTGGACTCACACGCGCTTCTCTTGGACTCTCGGACATCAAGATGGACCTGTTGAGTGGAT 1643
 Db 1546 TTGGATTTTGGGGTGTCTTAACCTGGCTTTTAAATAAAAGGAACAAGAAACT 1605
 QY 1644 GGAGAGAGTCTCTGCTGAAGGTCTCGCCGACATCAAGATGGACCTGTTGAGTGGAT 1703
 Db 1606 GGATCGCTTTTGGCTTCCAACTGTCCCAAGAGATAAGCAGCAAAATTCACAGTGCCT 1665
 QY 1704 CCAAGCAAAAGCTCAGAGCGAGCGCTCCACCTGCGAGGCGCTCCTTGGAGTTCTTCAG 1763

Db 1666 GAAGAGCTTAGGGAGCGTGGCAATCTCTCAGGAGCAGGTGGATTCCTTTGGCGATATTTTA 1725
Qy 1764 CTGCTTGTACAGATCAGAGGAGGAGTTTATCCAGCAGCGCTCTGAGCCACTTCCAGGT 1823
Db 1726 CTGCTCTTTGAATGAGAGTCTGCTTTGTGAGCGAGGAGTGAACCTTCTCCAG- 1784
Qy 1824 GATCGTGTACAGCAACTTGCCTCAAGATGAGCACAATGTTCTCTCTGTTCTGTCTGAA 1883
Db 1785 --AAGCTAACTTTCATATTTATGCAACAGTGGACTTGGTGGTTCTGCTACTGCTTAA 1842
Qy 1884 GCGTGTGAGGAGCGCCAGTGTGCTGATTTATGCGGCCACTACAGCGGGAGCGGA 1943
Db 1843 ATACTGCTCCAGCTTGAAGAACTCTGTTTTCGTTCCAAATGCTTTTAAAGAAAGAGA 1902
Qy 1944 AGACGCGGAGGTGCTCCGAGGAGCGCACAGCTGTTGTGAGCTGACGCTCAGACAGAGAG 2003
Db 1903 TGAACACAGCTCTAGTGGATTAAGCGCTCATCTGTTGGCATCATCTGCTCTGTTG- 1960
Qy 2004 GACCGTTCTGCTGAGCGCTACAGTGAACATCTGSCAGCGGCCCTGTGCAACAATCCAAA 2063
Db 1961 -----CTCACCAACAGCGGCA 1977
Qy 2064 CCGTATAGAGCTGTCTGTACCGAAATGCCCTGGCAGCGCGGGGTGAAGCTGCTGTG 2123
Db 1978 CCTCAGAGAGCTCCAGGTGCGAGGACAGCACCTCAGCGAGTGCACCTTTGTGACCTGTG 2037
Qy 2124 TCAGGAGCTCAGACACCCCAACTGCAAACTTCAGAACTCAGAGCTGGAAGGTGCGCAT 2183
Db 2038 TAACAGCTGAGGATCCAGCTGTGCTTCAGAGCTTGAATAAATACGTTTCCT 2097
Qy 2184 CTCAGCTCAGCGTGCAGGAGCTCTCTGAGCTCTCATAGCCAAATAGAAATTCGACAG 2243
Db 2098 TTCTGGCCAG--AGTGTCTCTCTTGTGAGGTGCTCTTTATCAGCAGACTTGAATA 2154
Qy 2244 GATGATCTCAGTGCACAGCGGTGGATTCAGCGATGATGCTGCTTTCGAGGGCT 2303
Db 2155 CCGTGTGCTACCTCCAGCAACTCTCTGATGACATCAGTCCCTCTGTGATGCTT 2214
Qy 2304 GCGGATCCAGTGCAGGCTCAGATGATTCAGTTGAGGAGTGTGAGTGCAGTGCAGTCCGG 2363
Db 2215 GAATACCCAGCAGCAGCTCAAGAGCTAGCGTGTGTAATGTCACCTCTCACCCT 2274
Qy 2364 GCGTGTGAGGAGTGGCTTCTGTGCTGGGCAACCAACCAATCTGTTGAGTTGAGCT 2423
Db 2275 TGATTTGAGTCTGCTGCTGCTGCTTCTAACCAACCAAGAGCTGAGTATCTGAATGT 2334
Qy 2424 GACGGAATGACTGGAGATTTGGGCTGAGTTTATCTATCCAGGAGCTGAGCACC 2483
Db 2335 ATCTGCAA--CCAGTTAGACACAGGCGTGCCTTTTGTGAGCGCTGTGAGCGCC 2391
Qy 2484 AGTCTGCAGACTACGAGCTTTTGGCTGAAGATCTGCGGCTCAGCTGCTGCTGTGA 2543
Db 2392 AGACAGGCTCTGTATACCTGATGTTGGCTTTCTGCCACCTCAGGAGCAGTGTGGA 2451
Qy 2544 CGAGTGGCTCACTCTCAGTGTGAACAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAA 2603
Db 2452 ATACATCTCTGAATGCTTCTGCTGAACAGAGCGTGCCTATCTAGACCTCAGTGCCAA 2511
Qy 2604 TGAGTGGGAGCTCGGGTCTGCTGTGTGAGGCGCTCAGGCTCCACCTGCA 2663
Db 2512 TGTCTGAAGAGCAGAGGACTGAAACTCTCTGCGAGGCTTGAACATCCGAGCTGTG 2571
Qy 2664 GCTCCAGAGCTCGGCTGGGATCTGCGGCTGGGCTCTGCGGCTGTGAGGCTTTTC 2723
Db 2572 CTGGATTCACTGTGTTTGTGAATGTTTATCATCTGCTGTGCTGTGAAGACCTGCG 2631
Qy 2724 TGTGTTGCTCCAGGCAACCAACCTCCGAGCTGAGTGTGAGTTTCAAGACCTGG 2783
Db 2632 CTCTGCTCTCATCAGCAATCAAACTGAAATCTCTGCAATTTGGTGTGAATGAAATCGG 2691
Qy 2784 AGACTGGGCGCTGTGCTGCTGAGGCGCTGCAACCTCCGCTGAGCTCCAGAA 2843
Db 2692 AGATGTGGTGTGACGCTGTTGTGTCGGCTCTGACGCAATCAGGATTCGCGCTAGAGAT 2751

Qy 2844 ACTGTGCTGATAGCTGTGTGCTCTCAGAGCAAGGCTTGTGAGAACTTTTACTTCCACCT 2903
Db 2752 TCTTGGTGTGAAGATGTGGTTTAAAGAGCAGCTGTGTAAAGATCTCGGCTGTGTCT 2811
Qy 2904 GGGATCAACAGAGCTTACCGACCTTACCTGACCAACAGCCCTAGGGGACACAGG 2963
Db 2812 CACCTGAGTAAAGCCCTGACAGCTCAACTGACCTTGAACACCTTGGACCAACAGG 2871
Qy 2964 TGTCCGACTGTCTTGAAGCGGCTGAGCCATCTCTGGCTGCAAACTCCAGTCTCTGTT 3023
Db 2872 GGTGTTCTACTCTGTGAGCCCTGAGACCCAGAGTGTCCCTGCGAGTGTCTGGCT 2931
Qy 3024 ATTTGGGATGACCTGAATAAATGACCCACAGTAGTGTGGAGCGCTTCAGTGAACAA 3083
Db 2932 GAGAAAACTGATTTGATGAGAAACCCAGGCATCTCTGACGCTGAGGAGAGAGAA 2991
Qy 3084 ACCTTATTGACAT 3098
Db 2992 TCCTAACCTGACCAT 3006

RESULT 13
BC012789
LOCUS BC012789.2 GI:33872518
DEFINITION Homo sapiens cDNA clone IMAGE:3957172, containing frame-shift errors.
ACCESSION BC012789
VERSION BC012789.2
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 3360)
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Woley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932
2 (bases 1 to 3360)
Strausberg, R.
Direct Submission

Submitted (15-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:15215377.

Contact: MGC help desk
Email: cgapps-r@mail.nih.gov

Tissue Procurement: ARCC
cdNA Library Preparation: Rubin Laboratory

cdNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>

REMARK
COMMENT

contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 25 Row: k Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 19745161
This clone has the following problem: frame shifted.

FEATURES

Location/Qualifiers
1..3360
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3957172"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH_MGC_21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

ORIGIN

Query Match 12.1%; Score 377; DB 11; Length 3360;
Best Local Similarity 48.9%; Pred. No. 1.7e-69;
Matches 1269; Conservative 0; Mismatches 1260; Indels 66; Gaps 7;
QY 513 CCTCTTTGACGACGAGGAGCGCCCGAGCGACCGCGCATGCTGCTGATGCAAGGCGC 572
DB CGCGCTTTTGTCTCCNAGGAACCTGGAAACAGCCACGTACAGTGAATATTCAAGGACC 528
QY 573 GGCAGGATAGGCAAGTCCATGCTGGCACACAAGGTTGATCTGAGCTGGCGCGACGGGAA 632
DB ACAAAGAAATGGAAAGACACTCTCTGATGAGCTGATGATGCGCTGGTGGCAACAA 598
QY 633 GCTCTTCCAGGAGGATTTGATTTATCTCTTCTATCACTCACTGAGGAGATGAACAGAG 692
DB GATCTTTTCGGGATAGGTTCTGTACAGTTCTATTTCTGTCGACAGAACTGAGGGAG-- 646
QY 693 TGGCAGGGAATGAGCATGCAAGACCTCATCTTTCAGCTGCTGCGCTGAGCCGACCGCGCC 752
DB TTTGCGCGCAACGAGTTTGGCTGACTTATTTCCAGAGAGTGGCCGTGACCCCGCTGCTCC 705
QY 753 TCTCCAGGAGCTCATCCGAGTTTCCGAGCGCTCTTTTCATCATCGACGGCTTCGATGA 812
DB TATAACAGAGATCGTGTCTCAACCGGAGAGACTCTTGTCTGTCATCGACAGCTTCGAGA 765
QY 813 GCTCAAGCTTCTTTCCAGATCTCTCAGGACCTCTGGTGCTCTGCTGGAGAGAAACG 872
DB GCTGCAAGGCGGCTTGAACGAAACCGATTGCGGATCTGTGTGTTGACTTGTATGGAAACG 825
QY 873 GCCCAGGAGCTGCTTCTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTT 932
DB GCGGTGCAAGTGCTTCTGAGCAGTTTGTGGAAGAGAGATGCTCCCGAGGCGCTCCCT 885
QY 933 GCTCAFCACCAACGCGCTTTTGGAGAGCTTCCACCGTCTGCTGAGCACCCCGAG 992
DB GCTCATCGCATCAAAACCGTGTGCCCGAAGAGCTCCGGATCAGTGAAGATCTCAGA 945
QY 993 GCATGGAGATCTCTGGGCTTCTGAGGAGAAAGAGGAATATCTTCAAGTATTT 1052
DB AATCTACCAAGCCCGGGATTCACGAGAGTATAGTGTATTTCTGCTGTTCTT 1005
QY 1053 CCACATGACAGACGCGGCGCAAGCTTCAATACGTGAGGAGCAACGAGGCTCTTT 1112
DB CAAGACCCGAAAGAGCCATGGAAGCTTCAATCTTGTGAAGAAAGTGAACAGCTGT 1065
QY 1113 CACCATGTGCTTCCCTCGTGGTGGTGGTGTGTACTGCTCCAGCAGAGCT 1172
DB TTCCATATGCCAAATCCCGCTCTCTGCTGGATCTGTGTACCAAGTCTGAAGCAAGAT 1125
QY 1173 GAGGGTGGGGGCTGTTGAGACAGACGTCAGGACCACTGAGTGTATGCTCTA 1232
DB GCAGAAAGGAAAGACCTGGCCCTGACCTGCCAGAGCACTACCTCTGTGTCTCTCTTT 1185

QY 1233 CCTGCTGAGTCTGATGCAACCCCAAG----CGGGGGCCCCGGGCTTCAGCCCCCACCACCA 1289
DB CGCTTTTAACTGTTTCACTGAGGTGCGGAGGCGCCGACTCCGCAAAACCCAGCACCA 1245
QY 1290 CCAGAGAGGCTTGTGCTTCTTGGCGGAGATGGGCTCTGGAATCAGAAAAATCTTATTGA 1349
DB GCTGAAGGCCCTGTGCTCTTGGCTGAGAGGATGTTGGACAGACATTTGAGTTTGG 1305
QY 1350 GGAGCGAGCTCCGGAAGCAACGCGCTTAGACGGGGAAGAGCTCTCTGCTTCTCAACAT 1409
DB TGAAGACGACTCCGGAGAAATGGGGTTGTGACGCTGACATCCCTGCGCTGCTGGGCAC 1365
QY 1410 GAACATCTTCCAGAGGACATCACTGTGAGAGTACTACAGCTTCACTCAGTTGAGTTT 1469
DB CAAGATACCTTCTGAAGTACGGGAGCGGTGAGAGCTCTACGTGTTCTCCACGCTGTAT 1425
QY 1470 CCAGGAATCTTTGACGCTATGTACTATATCTTGGACGAGGGGAGGCGGGCAGGCCC 1529
DB CAGGAGTCTGTGCGCGCTTGTCTATTGCTCAAGAGCCACTTGATCATCTCTCACC 1485
QY 1530 AGACCAAG-----GACGTGACCAAGGCTGTGACCGAGTACGCGTCTTCTGAAAGAGCTT 1583
DB AGCTGTGAGATGTGTACAGGAATTGCTGTGCAATTTTGAAGAAACAGGAGAGCACA 1545
QY 1584 CCTGGCACTCACCAGCGCTTCTGTTTGGATCTCTGAAAGAGAGACAGGAGCACCCT 1643
DB TTGGATTTTTTTGGGGTGTCTTAACTGGCTCTTTTAAATTAAGAAAGCAAGAAAACT 1605
QY 1644 GGAGAGAGTCTCTGTGGAAGTCTCGCGCACATCAAGATGACCTGTTGAGTGTGAT 1703
DB GGATGCTTTTGGGCTTCCACTGTCCAGAGATTAAGCAGCAATTCACAGTGCCT 1665
QY 1704 CCAAGCAAGCTCAGAGCGACGCGCTCCACCTGACAGCGGCTCTCTTGGAGTCTTTCAG 1763
DB GAAGAGCTTAGGGAGCGTGGCAATCTCAGGACAGGTGATCTCTTGGCGATATTTA 1725
QY 1764 CTGCTTGTACGAGATCCAGGAGGAGAGTTTATCAGCAGGCGCTGAGCCACTTCAGGT 1823
DB CTGCTCTTTTGAATACAGGATCTCTGCTTTGTAAGCAGGAGTGAACCTCTCCCAAG- 1784
QY 1824 GATCGTGTGAGCAACATTTGCTCCCAAGATGAGACATGCTCTCTCTGTTCTGTA 1883
DB --AAGCTAATTTTCAATTTATGACAACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1842
QY 1884 GGGCTCAGAGCGCCAGGCTGCTGCACTTTGATGGCGCCACCTACAGCGCGAGCGGGA 1943
DB ATACTGCTCCAGCTTCCAGGAACTCTGTTTTCGTTTCAAAATGTTCTTTAAGAAAGAGGA 1902
QY 1944 AGACCGCGGAGGTGCTCCGAGGAGCGCACGCTGTTGGTGGCTGAGCTCAGACAGAG 2003
DB TGAACACAGCTCTACGTCGGATTTACAGCCTCATCTGTTGGCATCATCTGCTCTGTG-- 1960
QY 2004 GACCGTCTGCTGGACGCTCAGTGAACATCTGGCAGCGGCGCTGTCACCAATCCAA 2063
DB -----CTCACCACAGCGGGCA 1977
QY 2064 CTGTATGAGCTTCTTCTGTAACGAAATGCCCTGGGAGCGCGGGGGTGAAGCTGCTCTG 2123
DB CCTCAGAGAGCTCCAGTTCAGGACAGACCCCTCAGCGAGTGCAGCTTTGTGACCTGGTG 2037
QY 2124 TCAAGACTCAGACACCCCAACTGCAAACTTCAAGAACCTGAGGCTGAGAGGTCGCGCAT 2183
DB TAAACAGCTGAGGATCCCGCTTCCAGAGCTTGGAAATTAATAAATGTTTCTT 2097
QY 2184 CTCAGCTCAGCTCGAGGAGCTCTGACGCTCTCATAGCCAAATAAGAAATTTGACAAG 2243
DB TTCTGGCCAG----AGTGTCTGCTCTTTGAGGTGCTCTTTATCAGCCAGACTTGAATA 2154
QY 2244 GATGATCTCAGTGGCAACGGGTTGGATTCCAGGCAATGATGCTCTTTTGGAGGCGCT 2303
DB CCTGAGCTTCACTTCAAGAACTCTCTGATGATGATGATGATGATGATGATGATGATGAT 2214

183 ATCAACACGACCTTTGACCGACCTTTTACCTGACCAACACGCGCTAGGGGACACAGGTGTC 242
2968 CGACTGCTTTGCAAGCGGCTGAGCCATCCTGGCTGCAAACTCCGAGTCTCTGGTTATTT 3027
243 CGACTGCTTTGCAAGCGGCTGAGCCATCCTGGCTGCAAACTCCGAGTCTCTGGTTATTT 302
3028 GGGATGGACCTGAATAAATAATGACCAAGTGGTTGGCAGCGCTTCGAGTACAAACCT 3087
303 GGGATGGACCTGAATAAATAATGACCAAGTGGTTGGCAGCGCTTCGAGTACAAACCT 362
3088 TATTTGGACATTGGCTGCTGA 3108
363 TATTTGGACATTGGCTGCTGA 383

Search completed: August 8, 2004, 07:04:46
Job time : 7585 secs

QY 2071 GAGCTGCTCTGTACCGAAATGCCCTGGGACGCGGGGG--TGAAGCTGCTGTCTCAAG 2128
Db 602 CTGTACTCTGTACCGAAATGCCCTGGGACGCGGGGGTGAAGCTGTGTCTCAAG 661
QY 2129 GACTCAGA 2136
Db 662 GACTCAGA 669

RESULT 15
BQ710940
LOCUS BQ710940
DEFINITION AGENCOURT 8487836 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6300641
5' mRNA sequence.
ACCESSION BQ710940
VERSION BQ710940.1 GI:21849839
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 983)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCM2515 row: d column: 18
High quality sequence start: 9
High quality sequence stop: 512.
Location/Qualifiers
1..983
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6300641"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

FEATURES
source
1..983
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6300641"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 113"
/note="Organ: spleen; Vector: pOTB7; Site 1: XhoI; Site 2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

ORIGIN
Query Match 11.4%; Score 354.6; DB 13; Length 983;
Best Local Similarity 97.4%; Pred. No. 4.8e-65;
Matches 371; Conservative 0; Mismatches 9; Indels 1; Gaps 1;
QY 2728 GTGCTCCAGGCCAACCAACCTCCGGGAGCTGGACTTGAAGTTTCAACGACCTGGGAGAC 2787
Db 4 GGGATTGGGACGAGGGCCCAACCTCCGGGAGCTGGACTTGAAGTTT-AACGACCTGGGAGAC 62
QY 2788 TGGGGCTGTGTTGCTGGCTGAGGGCTGCACATCCGCTCGAGATCCAGAAACTG 2847
Db 63 TGGGGCTGTGTTGCTGGCTGAGGGCTGCACATCCGCTCGAGATCCAGAAACTG 122
QY 2848 TGGCTGATAGCTGGGCTTCAGCCAGGGCTTGTGAGAACTTTTACTTACCCCTGGGG 2907
Db 123 TGGCTGATAGCTGGGCTTCAGCCAGGGCTTGTGAGAACTTTTACTTACCCCTGGGG 182
QY 2908 ATCAACACGACCTTGACCGACCTTTTACCTGACCAACACGCGCTAGGGGACACAGGTGTC 2967

PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 XX Stehlik C;
 DR WPI; 2002-471256/50.
 DR N-PSDB; ADE36416.
 XX
 PT Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 PS Claim 18; SEQ ID NO 24; 93pp; English.
 XX
 CC The invention describes an isolated PAAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NFkappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hyper trophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6.
 XX
 XX Sequence 1035 AA;
 Query Match 100.0%; Score 5472; DB 5; Length 1035;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MLRTAGDGLCRSLTYLEAEVLEAVELEKFKLYLGTATLGEKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGDGLCRSLTYLEAEVLEAVELEKFKLYLGTATLGEKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEANRLALSTPERINRKLWQREDELVRDQETRYDYVRKFRMLMEDRNARL 120
 DB 61 ITHFGPEANRLALSTPERINRKLWQREDELVRDQETRYDYVRKFRMLMEDRNARL 120
 QY 121 GECVNLSHRYTRLLLVKEHNPVQOQLDGTGRGHARTVGHQASPIKLTLPEDDEEP 180
 DB 121 GECVNLSHRYTRLLLVKEHNPVQOQLDGTGRGHARTVGHQASPIKLTLPEDDEEP 180
 QY 181 EPRTVMVQGAAGIGKSMALAHKVLMDWADGKLFQGRFDYLYFYNCRMNQSAECMSQDL 240
 DB 181 EPRTVMVQGAAGIGKSMALAHKVLMDWADGKLFQGRFDYLYFYNCRMNQSAECMSQDL 240
 QY 241 IFSCWPEPSAPLQELIRVPRLLFIIDGFDLKPSPHDPQPCWLCWEERKPTLLNSL 300
 DB 241 IFSCWPEPSAPLQELIRVPRLLFIIDGFDLKPSPHDPQPCWLCWEERKPTLLNSL 300
 QY 301 IRKLLPELSLLITRTPTALEKHLRLLEHPRHVEILGFSEAEKKEYFYKVFHNAEQGV 360
 DB 301 IRKLLPELSLLITRTPTALEKHLRLLEHPRHVEILGFSEAEKKEYFYKVFHNAEQGV 360
 QY 361 FNVVRDNEPLFTMCFVPLVCWVCTCLQQOEGGLLRQTSRTTAVMYLLYLSLMQPKP 420
 DB 361 FNVVRDNEPLFTMCFVPLVCWVCTCLQQOEGGLLRQTSRTTAVMYLLYLSLMQPKP 420
 QY 421 GAPRLQPPPNQGLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSFAFLNWNIFQKDINCE 480

DB 421 GAPRLQPPPNQGLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSFAFLNWNIFQKDINCE 480
 QY 481 RYYSFIHLSFOEFFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLATSRPLFGLL 540
 DB 481 RYYSFIHLSFOEFFAAMYIILDEGEAGAGPDQVTRLLTEYAFSERSFLATSRPLFGLL 540
 QY 541 NEETRSHLEKSLCWKVPKIMDLQWTSKAQSDGSTLQOGLSEFFSCLEYIOEEERFIQ 600
 DB 541 NEETRSHLEKSLCWKVPKIMDLQWTSKAQSDGSTLQOGLSEFFSCLEYIOEEERFIQ 600
 QY 601 QALSHFOVIVVSNIASRMEHVMVSSFCILKRCRQAQVHLHYGATYSADGEDRARCAGAHTL 660
 DB 601 QALSHFOVIVVSNIASRMEHVMVSSFCILKRCRQAQVHLHYGATYSADGEDRARCAGAHTL 660
 QY 661 LVQLRPRTVLLDAYSEHLAAALCTNPNLIELSYRNALSGRGVKLLCOGLRHPNCKLQ 720
 DB 661 LVQLRPRTVLLDAYSEHLAAALCTNPNLIELSYRNALSGRGVKLLCOGLRHPNCKLQ 720
 QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFPGVMLLCEGLRHPQCRLOMIQL 780
 DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFPGVMLLCEGLRHPQCRLOMIQL 780
 QY 781 RKQLESAGCAQEMASVLTGTPHVELDLTGNALEDGLRLCOGLRHPVCRRLTLMKIC 840
 DB 781 RKQLESAGCAQEMASVLTGTPHVELDLTGNALEDGLRLCOGLRHPVCRRLTLMKIC 840
 QY 841 RLTRAAACDELASTLVNQSLRELDLSNELDGLGVLLCEGLRHPTCKLQTLRLIGICRLG 900
 DB 841 RLTRAAACDELASTLVNQSLRELDLSNELDGLGVLLCEGLRHPTCKLQTLRLIGICRLG 900
 QY 901 SAACEGLSVVLOANHNRLDLSPNDLGMGLMGLLAEGLOHPACRLKWLDSGLTAKA 960
 DB 901 SAACEGLSVVLOANHNRLDLSPNDLGMGLMGLLAEGLOHPACRLKWLDSGLTAKA 960
 QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPCCKRLVWLFGMDLNKMTSHR 1020
 DB 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPCCKRLVWLFGMDLNKMTSHR 1020
 QY 1021 LAALRVTKPYLDIGC 1035
 DB 1021 LAALRVTKPYLDIGC 1035
 RESULT 2
 ID AAO15590 standard; protein; 1061 AA.
 AC AAO15590;
 XX
 DT 31-OCT-2002 (first entry)
 XX
 DE Human PYRIN-8 protein #2.
 XX
 KW Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-8.
 XX
 OS Homo sapiens.
 XX
 PN WO200261049-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002967.
 XX
 PR 31-JAN-2001; 2001US-0265231P.
 PR 10-SEP-2001; 2001US-0318645P.
 XX

PA	(MILL-) MILLENNIUM PHARM INC.	514	VTLLTEVAFSERSFLALTSRFLGLNNEETRSHEKSLCKWSPHIMKDWLQIQSKAQ	573
PA	(AMHP) WYETH.	541	VTLLTEVAFSERSFLALTSRFLGLNNEETRSHEKSLCKWSPHIMKDWLQIQSKAQ	600
XX				
PI	Bertin J, Wang W, Blatcher M;	574	SDGSTLQOGLSEFFSCLYEIQEEBFIQOALSHFQVIVVSNIAKMEHVMVSSFCLKRCRSA	633
XX				
DR	WPI; 2002-627477/67.	601	SDGSTLQOGLSEFFSCLYEIQEEBFIQOALSHFQVIVVSNIAKMEHVMVSSFCLKRCRSA	660
DR	N-PSDB; AAL44363.			
XX				
PT	New PYRIN polypeptides and nucleic acids useful for modulating and	634	QVHLHYGATYSADGEDRARCAGAHLLVQLRPERTVLLDAYSEHLAALCTNPNLIELS	693
PT	diagnosing stress-related, apoptotic and inflammatory responses, or for	661	QVHLHYGATYSADGEDRARCAGAHLLVQLRPERTVLLDAYSEHLAALCTNPNLIELS	719
PT	treating inflammatory and immune system disorders, cancers, or			
PT	neurological diseases.			
XX				
PS	Claim 8; Fig 8; 167pp; English.	694	LYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLITMDLSG	753
XX				
CC	The invention comprises the amino acid and coding sequences of human	720	LYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLITMDLSG	779
CC	PYRIN proteins. The PYRIN protein and DNA sequences of the invention are			
CC	useful for modulating and diagnosing stress-related, apoptotic and	754	NGVGPPGMMLICEGLRHPQCRLOMIQLRKCOLESGACQEMASVLGTNPHLVDELDTGNAL	813
CC	inflammatory responses. The PYRIN protein and DNA sequences are useful			
CC	for treating: inflammatory disorders and immune system disorders (e.g.	780	NGVGPPGMMLICEGLRHPQCRLOMIQLRKCOLESGACQEMASVLGTNPHLVDELDTGNAL	839
CC	Crohn's disease, reactive arthritis, multiple sclerosis, contact			
CC	dermatitis, psoriasis, graft rejection, allergies, viral infections and	814	EDLGLRLCOGLRHPVCRRLTLWLKICELTAACDELASTLSVNSOELSLNELGDL	873
CC	bacterial infections); cancer (e.g. leukemia); autoimmune disorders			
CC	(e.g. systemic lupus erythematosus and arthritis); and neurological	840	EDLGLRLCOGLRHPVCRRLTLWLKICELTAACDELASTLSVNSOELSLNELGDL	899
CC	diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN			
CC	protein and DNA sequences may also be used in screening assays, detection	874	GVLLICEGLRHPCTKQTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLW	933
CC	assays (e.g. chromosomal mapping, tissue typing or forensic biology),			
CC	predictive medicine (e.g. diagnostic assays, clinical trials and	900	GVLLICEGLRHPCTKQTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLW	959
CC	pharmacogenomics) and transcription profiling. The present amino acid			
CC	sequence represents a human PYRIN-8 protein	934	LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTTGINGINQTLTDLYLTNNALGDTGVRLIC	993
XX				
SQ	Sequence 1061 AA;	960	LLAEGLOHPACRLQKWLDSOGLTAKACENLYFTTGINGINQTLTDLYLTNNALGDTGVRLIC	1019
	Query Match 99.3%; Score 5433; DB 5; Length 1061;			
	Best Local Similarity 97.4%; Pred. No. 0;			
	Matches 1034; Conservative 0; Mismatches 0; Indels 28; Gaps 2;			
QY	1 MURTAGDGLCRSLSTYLELEAVELEKFKLYLGTATLGEGLKIPWGSMEKAGLEMAQLL	50		
Db	1 MURTAGDGLCRSLSTYLELEAVELEKFKLYLGTATLGEGLKIPWGSMEKAGLEMAQLL	60		
QY	61 ITHFGPEAWRLALSTFERINKDLWQREGLVLR-----	96		
Db	61 ITHFGPEAWRLALSTFERINKDLWQREGLVLR-----	120		
QY	97 ---DFQETRYDYRRKFRMLMEDNARLGEVNLSHRYTLLLVKHSNPMQVQQQLLDTG	153		
Db	121 PRKDPQETRYDYRRKFRMLMEDNARLGEVNLSHRYTLLLVKHSNPMQVQQQLLDTG	180		
QY	154 RGHARTVGHQAGPIKIETLFEDEERPEPRPTVMQAGIGKSMIAHKVMDWDAGKLF	213		
Db	181 RGHARTVGHQAGPIKIETLFEDEERPEPRPTVMQAGIGKSMIAHKVMDWDAGKLF	240		
QY	214 QGRFDYLFYINCREMMSQTECSMDLIQSCWPEPSAPLQELIRVPERLLFIIDGDELK	273		
Db	241 QGRFDYLFYINCREMMSQTECSMDLIQSCWPEPSAPLQELIRVPERLLFIIDGDELK	300		
QY	274 PSFHDPPQGPWCLEWEKRPTELLNSLRKLLPELSLLITRPTALEKHLRLEHPRHV	333		
Db	301 PSFHDPPQGPWCLEWEKRPTELLNSLRKLLPELSLLITRPTALEKHLRLEHPRHV	360		
QY	334 EILGFSEAEKKEYFYKFHNAEQAGQVFNVRNEPLFTMCFVPLVWVVCCLQOOLEG	393		
Db	361 EILGFSEAEKKEYFYKFHNAEQAGQVFNVRNEPLFTMCFVPLVWVVCCLQOOLEG	420		
QY	394 GGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEEDQ	453		
Db	421 GGLRQTSRTTAVMYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEEDQ	480		
QY	454 LRKHGLDGEDVSAFLNNMIFQKDINCERYYSFTHLSFQEFFAAMYIILDEGEGAGPDQD	513		
Db	481 LRKHGLDGEDVSAFLNNMIFQKDINCERYYSFTHLSFQEFFAAMYIILDEGEGAGPDQD	540		

RESULT 3	
AA017857	
ID	AA017857 standard; protein; 1099 AA.
XX	
AC	AA017857;
XX	
DT	20-AUG-2002 (first entry)
XX	
DE	Pyrin domain containing protein NALP3/PY5-hs.
XX	
KW	Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW	antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW	neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW	nephrotropic; osteopathic; neutropic; intracellular signal transduction;
KW	inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW	arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
XX	osteoarthritis; glomerulonephritis.
OS	Unidentified.
XX	
PN	WO200240668-A2.
XX	
PD	23-MAY-2002.
XX	
PF	30-OCT-2001; 2001WO-BP012545.
XX	
PR	15-NOV-2000; 2000DE-01056687.
PR	30-NOV-2000; 2000DE-01059595.
XX	
PA	(APOT-) APOTEC RES & DEV LTD.
XX	
PI	Tschoopp J, Martinon F;
XX	
DR	WPI; 2002-427093/45.
DR	N-PSDB; AAL47129.
XX	

PT New DNA encoding protein with pyrin domain, useful for treating diseases
 PT involving impaired signal transduction, particularly inflammation, also
 PT proteins and antibodies.

XX Claim 5; Fig 1; 116pp; German.

XX The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PVD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention

XX SQ Sequence 1099 AA;

Query Match	98.5%;	Score 5389;	DB 5;	Length 1099;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1019;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MLRTAGDGLCRISTYLELEAVELEAVELKPKYLGTATLGEKTPWGSMEYAGPLEMAQLL	60
Db	1	MLRTAGDGLCRISTYLELEAVELEAVELKPKYLGTATLGEKTPWGSMEYAGPLEMAQLL	60
QY	61	ITHFGPEEAWRLALSTFERINRNDLWERGOREDLVRDPQETRYDYVRKPERLMEDRNARL	120
Db	61	ITHFGPEEAWRLALSTFERINRNDLWERGOREDLVRDPQETRYDYVRKPERLMEDRNARL	120
QY	121	GECVNLSHRYTRILLVKEHSNPMQVQQQLDTRGSHARTVGHQASPIKIEFLPEPDEERP	180
Db	121	GECVNLSHRYTRILLVKEHSNPMQVQQQLDTRGSHARTVGHQASPIKIEFLPEPDEERP	180
QY	181	EPPTVVMQGAAGIGKSMIAHKVMDWADGKLFGQRPDYLYFYINCREMNSQATECSMDQL	240
Db	181	EPPTVVMQGAAGIGKSMIAHKVMDWADGKLFGQRPDYLYFYINCREMNSQATECSMDQL	240
QY	241	IFSCWPEPSAPLQELRVPERLFIIDGFDLKPSPHDPQWCLWEEKRPFELLNSL	300
Db	241	IFSCWPEPSAPLQELRVPERLFIIDGFDLKPSPHDPQWCLWEEKRPFELLNSL	300
QY	301	IRKKLLPELSLITTRPTALEKHLRLEHPRHVEILLGFSEARKEYFYKYFHNAEQAGQV	360
Db	301	IRKKLLPELSLITTRPTALEKHLRLEHPRHVEILLGFSEARKEYFYKYFHNAEQAGQV	360
QY	361	FNVVRNEPLFTWCFVPLVCWVCTCLOQLEGGLLRQTSRTTAVMYLLSLMQPKP	420
Db	361	FNVVRNEPLFTWCFVPLVCWVCTCLOQLEGGLLRQTSRTTAVMYLLSLMQPKP	420
QY	421	GAPRLQPPNORGLCSLAADGLWNQKILPEQDLRXHGLDGEDVSAFLNNNIFQKDINCE	480
Db	421	GAPRLQPPNORGLCSLAADGLWNQKILPEQDLRXHGLDGEDVSAFLNNNIFQKDINCE	480
QY	481	RYYSFIHLSFOEFAAMYIILDEGEAGDPQDVTRLLTEYAFSERSFLALTSRFLFGLL	540
Db	481	RYYSFIHLSFOEFAAMYIILDEGEAGDPQDVTRLLTEYAFSERSFLALTSRFLFGLL	540
QY	541	NEETRSHLEKSLCWKSPHVKMDLLOWIOSKASDGSSTLQOGSLFFSCLYETQEEFFIQ	600
Db	541	NEETRSHLEKSLCWKSPHVKMDLLOWIOSKASDGSSTLQOGSLFFSCLYETQEEFFIQ	600
QY	601	QALSHFQVIVVSNIAKMEHWSSFCILKRCRSQVLLHYGATYSADGEDRARCAGAHTL	660
Db	601	QALSHFQVIVVSNIAKMEHWSSFCILKRCRSQVLLHYGATYSADGEDRARCAGAHTL	660
QY	661	LVQLRPERTVLLDAYSEHIAAALCTPNLIELSLYNALGSRGVKLLCCGLRHPNCKLQN	720
Db	661	LVQLRPERTVLLDAYSEHIAAALCTPNLIELSLYNALGSRGVKLLCCGLRHPNCKLQN	720
QY	721	LRLKRCRISSACEDLSAALIANKNLITRMDLSGNGVGFPGWMLLCEGLRHPQCRLQMIOL	780
Db	721	LRLKRCRISSACEDLSAALIANKNLITRMDLSGNGVGFPGWMLLCEGLRHPQCRLQMIOL	780

QY	781	RKCOLBEGACQEWASVLGTPNPHLVELDLTGNALEDLGLRLCCGLRHPVCRRLTLWLKIC	840
Db	781	RKCOLBEGACQEWASVLGTPNPHLVELDLTGNALEDLGLRLCCGLRHPVCRRLTLWLKIC	840
QY	841	RLTAAACDELASTLSVNSQSLRELDLSINELGDLGVLLLCBGLRHPTCKLQTLRLGICRLG	900
Db	841	RLTAAACDELASTLSVNSQSLRELDLSINELGDLGVLLLCBGLRHPTCKLQTLRLGICRLG	900
QY	901	SAACEGLSVVLOQANHRLRELDLSFNLDGWLMLLAEGLOHPACRLOKWLDSGGLTAKA	960
Db	901	SAACEGLSVVLOQANHRLRELDLSFNLDGWLMLLAEGLOHPACRLOKWLDSGGLTAKA	960
QY	961	CENLYFTLGINQTLTDLYLTNALGDTGVRLLCCKRSHPCCKLRVLWLFQMDLKNKMTHS	1019
Db	961	CENLYFTLGINQTLTDLYLTNALGDTGVRLLCCKRSHPCCKLRVLWLFQMDLKNKMTHS	1019

RESULT 4	
ADG31287	
ID	ADG31287 standard; protein; 865 AA.
XX	
AC	ADG31287;
XX	
DT	18-DEC-2003 (first entry)
XX	
DE	Human novel polypeptide sequence, SEQ ID NO:1369.
XX	
KW	Human; diagnostic; drug screening; forensics; gene mapping;
KW	biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW	ulcers; osteoporosis; autoimmune disease; cancer;
KW	molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW	neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW	antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW	gene therapy; chromosome 19.
XX	
OS	Homo sapiens.
XX	
PN	WO2003029271-A2.
XX	
PD	10-APR-2003.
XX	
PF	24-SEP-2002; 2002WO-US030474.
XX	
PR	24-SEP-2001; 2001US-0324631P.
XX	
PA	(HYSE-) HYSEQ INC.
XX	
PI	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI	Haley-Vicente D, Drmanac RT;
XX	
DR	WPI; 2003-371981/35.
DR	N-PSDB; ADG30316.
XX	
PT	New polynucleotide and polypeptide useful for diagnosing, preventing or
PT	treating conditions such as neurodegenerative diseases, anemias, platelet
PT	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT	cancer.
XX	
PS	Claim 20; SEQ ID NO 1369; 1185pp; English.
XX	
CC	The invention relates to 971 novel human cDNA sequences (ADG29919-
CC	ADG30889) and the polypeptides they encode (ADG30890-ADG31860). The
CC	invention also relates to nucleic acid sequences over 99% identical with
CC	the novel human cDNAs. The invention additionally encompasses expression
CC	vectors and host cells comprising a nucleic acid of the invention; the
CC	recombinant production of a polypeptide of the invention; an antibody
CC	against a polypeptide of the invention; a method of detecting
CC	polynucleotides or polypeptides of the invention; and methods of
CC	identifying a compound which binds to a polypeptide of the invention. The
CC	invention further discloses methods of preventing, treating or

DR WPI; 2002-471256/50.
 XX Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX Claim 24; SEQ ID NO 64; 93pp; English.
 XX The invention describes an isolated PAAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical processes such as apoptosis, NFkappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6 leucine-rich-repeat domain (LLR).
 XX Sequence 603 AA;
 SQ

Query Match 57.8%; Score 3163; DB 5; Length 603;
 Best Local Similarity 100.0%; Pred. No. 8.9e-294;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNRGICSLAAGLWQKILFERODLRKHGLDGEDVSFAFLNNIFDKINCERYYSFIHL 488
 DB 1 PNRGICSLAAGLWQKILFERODLRKHGLDGEDVSFAFLNNIFDKINCERYYSFIHL 60
 QY 489 SFOEFAAMYIIDEGEGAGPDQDTRLLTEYAFSGRFLALTSRFLGLNEETRSHL 548
 DB 61 SFOEFAAMYIIDEGEGAGPDQDTRLLTEYAFSGRFLALTSRFLGLNEETRSHL 120
 QY 549 EKSLCWKVSPIHIMDLQWTSKAQSDGSTLQOQSLFFSCLYEIOEEFPIQALSHFQV 608
 DB 121 EKSLCWKVSPIHIMDLQWTSKAQSDGSTLQOQSLFFSCLYEIOEEFPIQALSHFQV 180
 QY 609 IIVSNITASKMEHVSFCLRCRCSAOLVLYGATYSADGEDRARCAGAHLLVQRPER 668
 DB 181 IIVSNITASKMEHVSFCLRCRCSAOLVLYGATYSADGEDRARCAGAHLLVQRPER 240
 QY 669 TVLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCOGLRHPCNKLQNLRLRCRI 728
 DB 241 TVLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCOGLRHPCNKLQNLRLRCRI 300
 QY 729 SSSACEDLSAALITANKNLTMDLSGNGVGPFGMMLLCEGLRHPCNKLQNLRLRCRI 788
 DB 301 SSSACEDLSAALITANKNLTMDLSGNGVGPFGMMLLCEGLRHPCNKLQNLRLRCRI 360
 QY 789 ACOEMASVLGTNPHLVELDITGNALDGLRLCCOGLRHPCNKLQNLRLRCRI 848
 DB 361 ACOEMASVLGTNPHLVELDITGNALDGLRLCCOGLRHPCNKLQNLRLRCRI 420
 QY 849 ELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCNKLQNLRLRCRI 908
 DB 421 ELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCNKLQNLRLRCRI 480
 QY 909 VVLQANHNLRELDLSFNDLGDWGLWLLAEGHQHPACRLQKLMWDSGLTAKACENLYFTL 968

DB 481 VVLQANHNLRELDLSFNDLGDWGLWLLAEGHQHPACRLQKLMWDSGLTAKACENLYFTL 540
 QY 969 GINQTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRVLLFGMDLNKMTSHRLAALRVTK 1028
 DB 541 GINQTLTDLVLTNNALGDTGVRLCKRLSHPGCKLRVLLFGMDLNKMTSHRLAALRVTK 600
 QY 1029 PVL 1031
 DB 601 PVL 603
 RESULT 6
 ABU99119
 ID ABU99119 standard; protein; 582 AA.
 AC ABU99119;
 XX
 DT 01-AUG-2003 (first entry)
 XX
 DE Novel human GPCR related protein NOV2a.
 XX
 KW Human; G-protein coupled receptor related protein; GPCR related protein;
 KW immunomodulator; cardiact; antiarteriosclerotic; antidiabetic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX
 OS Homo sapiens.
 XX
 PN WO200299116-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US017428.
 XX
 PR 04-JUN-2001; 2001US-0295607P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.
 PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.
 XX
 XX (CURA-) CURAGEN CORP.
 XX
 PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 PI Gangalli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;
 PI Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 PI Voss EZ, Zerhusen BD;
 WPI; 2003-140627/13.
 DR N-PSDB; ACD03623.
 XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.

PS Claim 1; Page 99; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27
CC 118-961 residue amino acid sequences, given in the specification, a
CC mature form of them, a sequence that is at least 95 % identical to them,
CC or a sequence having one or more conservative substitutions in them. The
CC polypeptide is useful in manufacturing a medicament for treating a
CC syndrome associated with a human disease selected from a pathology
CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
CC and antibodies are useful in treating or preventing NOVX-associated
CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
CC disease, Alzheimer's disease, infections; multiple sclerosis, cancer-
CC associated cachexia, and other wasting disorders associated with chronic
CC diseases. The nucleic acids and polypeptides may also be used as targets
CC for the identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This is the amino acid sequence of a novel human G-protein
CC coupled receptor related protein NOV

XX Sequence 582 AA;

Query Match 54.9%; Score 3005; DB 6; Length 582;
Best Local Similarity 99.3%; Pred. No. 1.2e-278;
Matches 578; Conservative 0; Mismatches 0; Indels 4; Gaps 3;

QY 373 MCFVPLVWVCTCLOQQLGGGLRQTSTRTTAVTMYLLSLMQPKGAPRLQPPNOR 432
DB 1 MCFVPLVWVCTCLOQQLGGGLRQTSTRTTAVTMYLLSLMQPKGAPRLQPPNOR 60
QY 433 GICSLAADGLWQKILFEQDLRKHGLEDGVSFLNMNIFOKDINCERYYSFIHLSFOE 492
DB 61 GICSLAADGLWQKILFEQDLRKHGLEDGVSFLNMNIFOKDINCERYYSFIHLSFOE 120
QY 493 FFAAMYIILDEGGAGPQDVTRLTAYAFSRSFALTSRFLGGLNNEETRSHEKSL 552
DB 121 FFAAMYIILDEGGAGPQDVTRLTAYAFSRSFALTSRFLGGLNNEETRSHEKSL 180
QY 553 CMKVSPIHIMDILLQWTSKAQSDGSTLQOQSLEFFSCLYEIOEEFIQALSHFQIVVS 612
DB 181 CMKVSPIHIMDILLQWTSKAQSDGSTLQOQSLEFFSCLYEIOEEFIQALSHFQIVVS 240
QY 613 NTASKMEHVVSSPCLKRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPERTVLL 672
DB 241 NTASKMEHVVSSPCLKRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPERTVLL 300
QY 673 DAYSEHLAALCTNPMLIELSYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRIS 731
DB 301 DAYSEHLAALCTNPMLIELSYRNALSGRGVKLLCOGLRHPNCKLQNLRLKRCRIS 360
QY 732 ACBDLSAALIANKNLFRMDLSGNGVFPGMMLICEGLRHQPCLQMLQRLKCOLESGACQ 791
DB 361 ACBDLSAALIANKNLFRMDLSGNGVFPGMMLICEGLRHQPCLQMLQRLKCOLESGACQ 420
QY 792 EMASVLGTNPHLVDELDTGNALEDGLRLCCOGLRHPVCRRLTLW-LKICRLTAACDE 849
DB 421 EMASVLGTNPHLVDELDTGNALEDGLRLCCOGLRHPVCRRLTLW-LKICRLTAACDE 480
QY 850 LASTLSVNQSLRELDLSINELGDLVLLCEGLRHPTCKLQTLRLGICRLGSAACBGLS 908
DB 481 LASTLSVNQSLRELDLSINELGDLVLLCEGLRHPTCKLQTLRLGICRLGSAACBGLS 540
QY 909 VVLQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKLM 950

DB 541 VVLQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKLM 582

RESULT 7

ABG97475

XX ABG97475 standard; protein; 565 AA.

XX ABG97475;

XX 16-DEC-2002 (first entry)

XX Human nucleic acid associated protein, NAAP9, from, INCYTE no.429930CD1.

XX Human; nucleic acid associated protein; NAAP; cancer;
KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
KW atopic dermatitis; arthritis; bacterial infection; viral infection;
KW parasitic infection; protozoal infection; fungal infection.

XX Homo sapiens.

XX WO200272630-A2.

XX 19-SEP-2002.

XX 07-FEB-2002; 2002WO-US003844.

XX 09-FEB-2001; 2001US-0268118P.

XX 21-FEB-2001; 2001US-0270963P.

XX 22-FEB-2001; 2001US-0270858P.

XX 23-FEB-2001; 2001US-0271194P.

XX 07-MAR-2001; 2001US-0284071P.

XX 12-APR-2001; 2001US-0283496P.

XX 09-NOV-2001; 2001US-0344650P.

XX (INCY-) INCYTE GENOMICS INC.

XX Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Wallia NK, Thangavelu K;
PI Elliott VS, Marquis JP;

XX WPI; 2002-723320/78.

XX N-PSDB; ABS78719.

XX New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.

XX Claim 1; Page 147-149; 162pp; English.

XX The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
CC screening for a compound that is effective as an ant/agonist or modulator
CC of NAAP, generating an expression profile of a sample containing the
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridisable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders

CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
 CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
 CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
 CC other diseases and disorders listed in the specification. These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of NAAp. The NAAp or its
 CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence represents an NAAp protein
 CC
 XX SQ Sequence 565 AA;

Query Match 54.0%; Score 2953.5; DB 5; Length 565;
 Best Local Similarity 99.8%; Pred. No. 1e-273;
 Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	470	MMIPQKDCINERYYSFTHLSQFEEFAAMYYILDEGGAGDPDQVTLTTEYAFSESL	529
Db	1	MMIPQKDCINERYYSFTHLSQFEEFAAMYYILDEGGAGDPDQVTLTTEYAFSESL	60
QY	530	ALTSRFLGLLNEETSHLEKSLCKWYSPIKMDLLQWIOSKQSDGSTLQOQSLEFFSC	589
Db	61	ALTSRFLGLLNEETSHLEKSLCKWYSPIKMDLLQWIOSKQSDGSTLQOQSLEFFSC	120
QY	590	LYIQEIEEFTQALSHFQVIVVNIASKMEHMYSSFCFKCRSAQVHLHYGATYSADGED	649
Db	121	LYIQEIEEFTQALSHFQVIVVNIASKMEHMYSSFCFKCRSAQVHLHYGATYSADGED	180
QY	650	RARCSAGHTLLVQLPERVTLDAVSEHLAALCTNPNIETSLVENALGSGVKKLQ	709
Db	181	RARCSAGHTLLVQLPERVTLDAVSEHLAALCTNPNIETSLVENALGSGVKKLQ	239
QY	710	GLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR	769
Db	240	GLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR	299
QY	770	HPQCRLOMIQRLKQLESGAQEMASVLGNPHVLDLTGNALDGLRLCCGLRHPV	829
Db	300	HPQCRLOMIQRLKQLESGAQEMASVLGNPHVLDLTGNALDGLRLCCGLRHPV	359
QY	830	CRRLTWLTKICRTAAACDESLASTLSVNQSLRELDLSLNELDGLVLLCEGLRHPTCKL	889
Db	360	CRRLTWLTKICRTAAACDESLASTLSVNQSLRELDLSLNELDGLVLLCEGLRHPTCKL	419
QY	890	QTLRLGICRIGSAACGLSVVLQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKL	949
Db	420	QTLRLGICRIGSAACGLSVVLQANHNRLDLSFNDLGDWGLWLLAEGLOHPACRLQKL	479
QY	950	WLDSCGLTAKACENLYFTGINTDLYTNALGDTGVRLICKRLSHPGCKLRVLWLP	1009
Db	480	WLDSCGLTAKACENLYFTGINTDLYTNALGDTGVRLICKRLSHPGCKLRVLWLP	539
QY	1010	GMDLNKMTSHSLAALRVTPYLDIGC 1035	
Db	540	GMDLNKMTSHSLAALRVTPYLDIGC 565	

RESULT 8

ABU99120

ID ABU99120 standard; protein; 521 AA.

XX AC ABU99120;

XX AC ABU99120;

XX DT 01-AUG-2003 (first entry)

XX DE Novel human GPCR related protein NOV2b.

XX DE Human; G-protein coupled receptor related protein; GPCR related protein;

XX KW

KW NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 XX
 XX Homo sapiens.
 OS
 XX WO200299116-A2.
 PN
 XX
 PD 12-DEC-2002.

04-JUN-2002; 2002WO-US017428.

04-JUN-2001; 2001US-0295607P.

04-JUN-2001; 2001US-0295661P.

06-JUN-2001; 2001US-0296404P.

06-JUN-2001; 2001US-0296418P.

14-JUN-2001; 2001US-0298285P.

15-JUN-2001; 2001US-0298556P.

21-JUN-2001; 2001US-0299949P.

26-JUN-2001; 2001US-0300883P.

28-JUN-2001; 2001US-0301550P.

13-AUG-2001; 2001US-0311972P.

27-AUG-2001; 2001US-0315071P.

29-AUG-2001; 2001US-0315660P.

14-SEP-2001; 2001US-0322293P.

17-SEP-2001; 2001US-0322706P.

14-DEC-2001; 2001US-0341186P.

28-FEB-2002; 2002US-0361189P.

12-MAR-2002; 2002US-0363673P.

12-MAR-2002; 2002US-0363676P.

03-JUN-2002; 2002US-00363676.

(CURA-) CURAGEN CORP.

Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 Gangoli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 MacDougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;
 Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 Voss EZ, Zerhusen BD;

WPI; 2003-140627/13.

N-PSDB; ACD03624.

New NOVX polypeptides and nucleic acids, useful for preventing or
 treating NOVX-associated disorders e.g. cancer, cardiomyopathy,
 atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 pharmacogenomics.

Claim 1; Page 100; 332pp; English.

The invention describes an isolated polypeptide (I) comprising any of 27
 118-961 residue amino acid sequences, given in the specification, a
 mature form of them, a sequence that is at least 95 % identical to them,
 or a sequence having one or more conservative substitutions in them. The
 polypeptide is useful in manufacturing a medicament for treating a
 syndrome associated with a human disease selected from a pathology
 associated with the polypeptide. The NOVX polypeptides, polynucleotides
 and antibodies are useful in treating or preventing NOVX-associated
 disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 associated cachexia, and other wasting disorders associated with chronic
 diseases. The nucleic acids and polypeptides may also be used as targets
 for the identification of small molecules that modulate or inhibit e.g.
 neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 wound healing and angiogenesis, in gene therapy, in generation of

QY 561 KMDLLQWIOKQSDGSLTQOGLRFFSCLYEIQEERTQOALSDFQVIVVSNIAKMEH 620
 Db 603 RLELKWLEVKAKKQIQPSQLSELYCYEQBEDFVQRAWDYFPKTEI-NLSTRMDH 661
 QY 621 MVSSFLKRCRAQVILHLYGAYSDGEDRCASAGATHLLVQLRPETVILDAYSEHLA 680
 Db 662 MVSSFCIENHVESLSL-GFLHNPKEEKEEKEGRHLDWQ----- 703
 QY 681 AALCTNPNIELIYRNALSGVKKLCOGLRHPNCKLQNLKRCRISSSACEDLSAAL 740
 Db 704 ---CVLPS-----SSHAACSHG-----LVNSHLTSSFCRGLFSLV 735
 QY 741 IANKNLTMDLSNGVGPFGWMLCEGLRHPQCRLOMIOLRKOLESAGCQEWASVLGTN 800
 Db 736 STSQSLTELDLSNLSGDPGMVLCETLQHPGNCIRRLMIGRCSHCCEFDLSLVSSN 795
 QY 801 PHLVDELDTGNALDELGLRLCQGLRHPVCRILTTLWTKICRTAAACDELASTLSVNOSL 860
 Db 796 QKLVDELSDNALGDFGIRLLCVGLKHLNCKLWLVSSCCLTSACCCQDLASVLSHSL 855
 QY 861 RELDLSNELGDLGVLLCEGLRHPTCKLQTLRLGICRIGSACGLSVVLQANHLREL 920
 Db 856 TRLYVGENALGDSGVAILECAKNPCQNLQKGLVNSGLTSVCCSALSLSVLTNQNLTHL 915
 QY 921 DLSFNDLGDWGLWLLAEGLOHPACRLOKLPDSCGLTAKACENLYFTLGINOTLTLXYLT 980
 Db 916 YURGMTLGDGKHLKCEGLLHPDCKLQVLELONCNITSCCWDLSLTLTSQSRLKLSLG 975
 QY 981 NNALGDTGVRLLCRISHPGCKLRVLWLFPMGLNKNMTHSRLAALRVTKPYLDI 1033
 Db 976 NNDLGLDGLWMMCEVLKQOQSLQLNGLSEMYFNVTYSKSALETLOEKEPALTIV 1028
 RESULT 10
 ABU08503
 ID ABU08503 standard; protein; 1034 AA.
 AC ABU08503;
 XX
 XX
 DT 22-MAY-2003 (first entry)
 DE Huma PYRIN-1 protein.
 XX Human; nucleotide binding site; pyrin domain; NBS-1; PYRIN-1; caspase-1;
 KW antiinflammatory; apoptosis; ASC; NF-KB; nuclear factor KB; IRR;
 KW leucine rich repeat; inflammatory disorder; familial cold urticaria;
 KW arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 KW ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 KW insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 KW psoriasis; graft versus host disease; food allergy; conjunctivitis;
 KW chronic obstructive pulmonary disease; food allergy;
 KW helminthic infection; leishmaniasis; viral infection; HIV infection;
 KW bacterial infection; tuberculosis; leprosy; chromosome 1q44.
 XX
 OS Homo sapiens.
 XX
 XX
 FN US2002187922-A1.
 XX
 PD 12-DEC-2002.
 XX
 PF 22-APR-2002; 2002US-00127516.
 XX
 PR 17-FEB-2000; 2000US-00506067.
 PR 01-SEP-2000; 2000US-00653901.
 PR 26-SEP-2001; 2001US-00964955.
 PR 20-DEC-2001; 2001US-00027629.
 XX
 XX (BERT/) BERTIN J.
 PA (MANJ/) MANJI G A.
 XX
 PI Bertin J, Manji GA;
 XX
 XX WPI; 2003-328763/31.
 DR

DR N-PSDB; ABX93556.
 XX Identification of compound that binds to polypeptide (for e.g. a PYRIN protein), useful for treating disorders associated with inappropriate apoptosis, for e.g. inflammatory disorder.
 PT Claim 1; Fig 4; 74pp; English.
 CC The invention relates to a new method for the identification of a compound that binds to a human PYRIN-1 (an apoptotic signalling molecule appearing as ABU08503) comprising contacting the polypeptide or a cell expressing the polypeptide to a test compound and determining whether PYRIN-1 binds to the test compound, or determining the effect of the test compound on the activity of PYRIN-1. In a similar manner, compounds are isolated which modulate the binding of PYRIN-1 to ASC (comprising testing the compounds against the pyrin binding domains of both PYRIN-1 and ASC (not defined)), modulate the activity of NF-kB (nuclear factor KB), and which modulate the ASC-mediated activation of NF-kB (comprising: (a) measuring the binding of a test compound to the LRR (leucine rich repeat) domain of PYRIN-1; and (b) measuring the activation of NF-kB in a cell expressing ASC and PYRIN-1 in the presence and absence of the LRR domain binding compound). The candidate modulators may be useful for treating an inflammatory disorder. Also included is a method (M8) for identifying a modulator of caspase-1 activity. The identified compound can be used to modulate the polypeptide's activity (ASC and NF-kB activities in a patient). By modulating the expression or activity of the polypeptide (PYRIN-1 and ASC), a disorder associated with inappropriate apoptosis (inflammatory disorders e.g. familial cold urticaria, arthritis, inflammatory bowel disease, Crohn's disease, ulcerative colitis, rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus host disease, asthma, chronic obstructive pulmonary disease, allergies (e.g. food allergies), conjunctivitis, helminthic infection (e.g. leishmaniasis), viral infections such as HIV infection, and bacterial infections such as tuberculosis and leprosy) can be treated. Human NBS-1 protein (nucleotide binding site protein 1) contains a pyrin domain which was used to identify genomic sequences containing the PYRIN-1 gene. The gene for pyrin-1 is located on chromosome 1q44. The present sequence represents human PYRIN-1

Sequence 1034 AA;

Query Match 43.8%; Score 2396; DB 6; Length 1034;
 Best Local Similarity 46.1%; Pred. No. 8.5e-220;
 Matches 495; Conservative 175; Mismatches 303; Indels 100; Gaps 10;
 QY 11 CRSLTYLELEAVELEKFKLYL-GTATELGEKGIPIWGSMEKAGPLEMAQLLTHFGPEPA 69
 Db 6 CKLARYLEDLEDVLDLKKFKWHLDDYPPQKGCIPLRGQTEKADHDVLAFLMIDFNGEKA 65
 QY 70 WRLALSTTERINRKLWREGQRED----- 93
 Db 66 WAMAVWIFAAINRRDLYEKAKDEPKWGSNDARVSNPTVICOEDSIEEWMGLLEYLSRI 125
 QY 94 ----LVDPDQETRYDVRKPRLMEDRNARLCECNLSHRYTRLLLVKHSNPMQVQOOL 149
 Db 126 STCKMKKYRKYRYKYSRFOCIEDRNARLGESVSLNRYTRLRILKEHRSQREQL 185
 QY 150 LDTGRGHARTVGHQASPIKXETLFEPEDEPERPPRTVYVQGAAGIKSMIAHKVMDWAD 209
 Db 186 LAIGK-TKTCSPVSPKIMKELLFDPDDEHSEPVHTVVFQGAAGIKTILARKWMDWAS 243
 QY 210 GKLFQGRFDYLYINCREMNQSAATSCMDLIFSCWPEPSAPLOBLIRVPELLFIIDGF 269
 Db 244 GTLYQDRFDYLYIHCREVS-LVTQSLGDLIMSCCPDPNPPIHKIVRKPSILFLMDGF 302
 QY 270 DELKPSFHDPPQGPWCLCWEEKRPTLLNLSLRKLLPELSLITTRPALCKHLRLH 329
 Db 303 DELQAFDEHIGLPLCTDQWKAERGDIILSLLRKLLPEASLITTRPALEKILQLLDH 362
 QY 330 PRHVEILGFSEAKREYFKYFHNAAEQAGQVFNVRDNEPFTMCFVPLVCVWCTCLOQ 389
 Db 363 PRHVEILGFSEAKREYFKYFHNAAEQAGQVFNVRDNEPFTMCFVPLVCVWCTCLOQ 422

QY	390	OLEGGGLQTSRTTAVYMLYLLSLMQPKAPRQPPNQRGLCSLAADGLWNOKILF	449	FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
FT		164. .167		FT		/note= "CAMP/cGMP-dependent protein kinase phosphorylation site"
Db	423	QMESGKSLAQTSTKTAVYVFLSSLLQPRGSGEHLCAHLWGLCSLAADGLWNOKILF	482	FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
QY	450	EBQDLRKHGLDGEDVSAFLNMNIFQDINCERYYSFIHLSFQEFFAAMYIILDEGGGAG	509	FT	Domain	/label= "Expanded nucleotide binding site_domain"
Db	483	EBSDLRNHGLQKADVSAFLRMNLFOKEVDCERFYSFIHMTFQEFFAAMYIILDEEGRT	542	FT	Region	/note= "NACHT NTPase"
QY	510	-----PDQVTLLEYAFSERSFLALTSRFLFGLLNETHRSHLEKSLCWKVSPIH	560	FT	Modified-site	/label= "Expanded nucleotide binding site domain signature motif #1"
Db	543	NVPGSRKLPSRDVTLLNENYKFERGYLFFVVRFLGLVNOBRTSYLEKLSCKISQOI	602	FT	Domain	/label= "Kinase Ia_domain"
QY	561	KMDLLWIOSKAQSDGTLTQOGSEFFSCLYETQEEFIQALSHFQVIVWSNIASKMEH	620	FT	Region	/note= "ATP/GTP-binding site motif A (P-loop)"
Db	603	RUELLKWEIVAKAKKLOQPSQLELFYCLYEMQEDFVORANDYFPKTEI-NLSTRMDH	661	FT	Modified-site	/label= N-myristoylation_site
QY	621	MVSSFCLKRCRAQVILHYGATYSADGEDRARCAGAHITLLVQLRPERTVLLDAYSEHLA	680	FT	Region	/note= "Expanded nucleotide binding site domain signature motif #2"
Db	662	MVSSFCIENCHRVESLSL-GFLHNPKEEBEKEGRHLDVQ-----	703	FT	Domain	/label= Nucelotide_binding_site_domain
QY	681	ALCTPNLIELSLYNALGSRGVKLLCOGLRPNCKLQNLKRCRISSSACEDLSAAL	740	FT	Modified-site	/label= Protein_kinase_C_phosphorylation_site
Db	704	---CVLP-----SSHAACSHG-----LVNSHLTSSFCRGLFSVL	735	FT	Modified-site	/label= Casein_kinase_II_phosphorylation_site
QY	741	IANKNLTRMDLSNGVGFPGMMLLCEGLRHPOCRLOMIQLRKCQLESAGQEMASVILGTN	800	FT	Region	/note= "Walker B_box"
Db	736	STQSITELDSLSNSIGDPGRVLCVTLQHPGNIIRLMLGRGLSHECCFDLSVLSSN	795	FT	Domain	/label= Kinase 2 domain
QY	801	PHIVELDLTGNALDELGLRLCOGLRHPVCRRLTLMKICRLTAAACDELASTLVNQSL	860	FT	Modified-site	/note= "CAMP/cGMP-dependent protein kinase phosphorylation site"
Db	796	QKVELDLSNALGDFGIRLLCVGLXHLNLNKKMLVSCCLTSACQDLASVLSTSHSL	855	FT	Region	/note= "RGD cell attachment sequence"
QY	861	RELDLSNELGLDGLVLLCEGLRHPTCKLTBLGICRLGSAACEGLSVLVQANHLREL	920	FT	Domain	/label= Protein_kinase_C_phosphorylation_site
Db	856	TRLVYGENALGDSGVAILECAKPNQCNLOKGLVNSGLTSVCCSALSSVLTQNQLTHL	915	FT	Modified-site	/note= "Expanded nucleotide binding site domain signature motif #5"
QY	921	DLSFNDLGDWGLWLLAAGLQHPACRLQKWLDSGLTAKACENLYFTLGINQTITDLYLT	980	FT	Domain	/label= Kinase_3a_domain
Db	916	YLRGNVLGDKGKILCEGLLHPCKQLVLELDNCLTSHCCWDLSTLTSSQSLKLSLG	975	FT	Modified-site	/label= Protein_kinase_C_phosphorylation_site
QY	981	NNALGDTGVRELLCKRLSHPCCKLRVLWLFQMDLNKMTSHSLAALRVTKPYLDI	1033	FT	Region	/note= "Dileucine motif"
Db	976	NNDLGLGVNMFCEVLKQSQCLLQNLGLSEMYFNETKSALETLOEKPELTV	1028	FT	Modified-site	/label= N-myristoylation_site
RESULT 11						
ID	ABU63315					
XX	AC	ABU63315 standard; protein; 1034 AA.				
XX	AC	ABU63315;				
XX	XX	23-SEP-2003 (first entry)				
DE	DE	Human pyrin domain family protein PYRIN-1.				
KW	KW	Human; pyrin domain family; PYRIN-1; cellular differentiation; NF-kappaB;				
KW	KW	tissue typing; cellular proliferation; cell survival; apoptosis disorder;				
KW	KW	inflammatory disorder; apoptosis associated speck like protein; ASC;				
XX	XX	nuclear factor kappaB.				
OS	OS	Homo sapiens.				
PH	PH	Key	Location/Qualifiers			
FT	Domain	1. .87	/label= Pyrin_domain			
FT	Modified-site	3. .5	/label= Protein_kinase_C_phosphorylation_site			
FT	Modified-site	44. .46	/label= Protein_kinase_C_phosphorylation_site			
FT	Modified-site	93. .98	/label= N-myristoylation_site			
FT	Modified-site	110. .113				

phosphorylation site"
 595. .597
 /label= Protein_kinase_C_phosphorylation_site
 606. .607
 /note= "Dileucine motif"
 618. .626
 /note= "peroxisomal targeting sequence"
 624. .627
 /label= Casein_kinase_II_phosphorylation_site
 654. .657
 /note= "Asn is glycosylated"
 656. .658
 /label= Protein_kinase_C_phosphorylation_site
 657. .660
 /label= Casein_kinase_II_phosphorylation_site
 717. .722
 /label= N-myristoylation_site
 740. .991
 /label= Leucine_rich_repeat_domain
 740. .767
 /label= Leucine_rich_repeat
 740. .743
 /label= Casein_kinase_II_phosphorylation_site
 750. .753
 /label= Casein_kinase_II_phosphorylation_site
 769. .796
 /label= Leucine_rich_repeat
 797. .821
 /label= Leucine_rich_repeat
 815. .816
 /note= "Dileucine motif"
 816. .837
 /note= "Leucine zipper pattern"
 823. .824
 /note= "Dileucine motif"
 826. .849
 /label= Leucine_rich_repeat
 854. .878
 /label= Leucine_rich_repeat
 883. .906
 /label= Leucine_rich_repeat
 888. .893
 /label= N-myristoylation_site
 911. .935
 /label= Leucine_rich_repeat
 911. .914
 /note= "Asn is glycosylated"
 919. .924
 /label= N-myristoylation_site
 921. .924
 /label= Casein_kinase_II_phosphorylation_site
 929. .930
 /note= "Dileucine motif"
 934. .935
 /note= "Dileucine motif"
 940. .967
 /label= Leucine_rich_repeat
 950. .953
 /note= "Asn is glycosylated"
 962. .963
 /note= "Dileucine motif"
 988. .991
 /label= Leucine_rich_repeat
 968. .970
 /label= Protein_kinase_C_phosphorylation_site
 970. .973
 /note= "cAMP/cGMP-dependent protein kinase phosphorylation site"
 997. .998
 /note= "Dileucine motif"
 1014. .1017
 /label= Casein_kinase_II_phosphorylation_site
 1018. .1021

/label= Casein_kinase_II_phosphorylation_site
 US2002197660-A1.
 26-DEC-2002.
 20-DEC-2001; 2001US-00027629.
 17-FEB-2000; 2000US-00506067.
 01-SEP-2000; 2000US-00653901.
 26-SEP-2001; 2001US-00964955.
 (BERT/) BERTIN J.
 (MANJ/) MANJI G A.
 Bertin J, Manji GA;
 WPI; 2003-521543/49.
 N-PSDB; ACD27909.
 Identifying compound that binds to polypeptide useful in treating
 PT apoptotic and inflammatory disorders, comprises contacting polypeptide
 PT with test compound and determining whether polypeptide binds to test
 PT compound.
 XX
 PS Claim 1; Fig 4; 71pp; English.
 XX
 CC The invention relates to a compound that binds to a polypeptide (e.g.
 CC PYRIN-1 or NBS-1). The nucleic acid molecules, proteins, protein
 Query Match 43.8%; Score 2396; DB 6; Length 1034;
 Best Local Similarity 46.1%; Pred. No. 8.5e-220; Indels 100; Gaps 10;
 Matches 495; Conservative 175; Mismatches 303;
 QY 11 CRLSTYLEEAEVLEKFKLYL-GTATELGEGKIPWGMKAGPLEMAQLLTHFGPEA 69
 Db 6 CKLARYLEDLEDVDLKKFKHLEDYPPQKGCPLPRGQTERKADHVDLATLMDFNGEKA 65
 QY 70 WRLALSTFERINRDLWEGQRED----- 93
 Db 66 WAMAVWIFAANRRDLVEKAKRDEPKWGSNDARVSNPTVICQEDSIEEWMGLBYLSRI 125
 QY 94 ----LVROPQETRYDYVRKFLMEDNRALGECVNLSHRYTRLLLVKHSNPMOVQOOL 149
 Db 126 SICWKDYRKRYKRYVRSRFOCIEDNRALGESVSNKRYTRLRILKEHRSQREQL 185
 QY 150 LDTGRGHARTVGHQASPIKIELTFEPDEERPPRTVVMQGAAGICKSLAHKVMLDWAD 209
 Db 186 LAIGK--TKTCSPVSPKIMELLFPDDDEHSEPVHTVVFQGAAGIKTLARKMMLDWS 243
 QY 210 GKLFQGRFDYLYINCRMNQSAECMSQDILFSCWBPSPALQELIRVPERLLFIIDGF 269
 Db 244 GTLYQDRFDYLYIHCREVS-LVTQRLGDLIMSCCPDPNPIHKIVRKPSRILFLMDGF 302
 QY 270 DELKPSFHDPOGPMCLCWEEKRPTELLNLSLRKLLPELSLLITRPTALEKHLRLLEH 329
 Db 303 DELQAFDEHIGPLCTDQKAEGRDILLSLRKLLPEASLLITRPTALEKHLRLLEH 362
 QY 330 PRHVEILGFSEAKREYFYKYPHNAEQAGQVFNVDNEPLFTMCFVPLVCVWCTCLOQ 389
 Db 363 PRHVEILGFSEAKREYFYKYPHNAEQAGQVFNVDNEPLFTMCFVPLVCVWCTCLOQ 422
 QY 390 QLEGGGLRQTSRTTAVTMYLLSLIMQPKCAPLQPPPNQGRGLCSLAADGLWQKILF 449
 Db 423 QMESGKSLAQTSKTTTAVYVFFLSLLQPRGSGQEHGLCAHLWGLCSLAADGLWQKILF 482
 QY 450 EEDLRKGLDGEDVSFAFLNMNI FOKDINCERYYSFIHLSFOEFAAMYIILDEBEGGAG 509
 Db 483 EESDLRNHGLQKADVSAPFLRMNLFQKEDVCFKFIEMTFQEFFAAMYILLEEKGT 542
 QY 510 -----PDQDVTLLTEYAFSERSFLALTSRFLGLNNEETRSHLEKSLCWKVSPIH 560
 Db 543 NYPGSRKLPSRDVTVLLLENYKFEKGYLIFVVRFLFGLVQERTSYLEKSLCKISQOI 602

QY 561 KMDLLQWIOSKAGSDGSLTQOGLRPFSCLYEIOEEFIQOALSHFOVIVWSNIASKEH 620
 Db 603 RLELLKWIIEKAKAKLQPSOLELFCYLYEIOEEFVQRAWDPFKIEI-NLSTRMDH 661
 QY 621 MVSSFCIKRCRGAQVHLHYGATYSDGDEDRARCSAGAHTLLVQLRPERVTLDAYSEHLA 680
 Db 662 MVSSFCIENCHRVEISL-GFLHNMPEKEEKEGRHLDWQ----- 703
 QY 681 ALCTNPNIIELSLYNALSRSQVKKLLCOGLRHPNCKLQNLKRCRISSACEDLSAAL 740
 Db 704 ---CVLPS-----SSHAACSHG-----LVNSHLTSSFCRGLFSVL 735
 QY 741 IANKNITRMDLSGNGVGFQGMMLLCEGLRHPQCRLOMIOLRKOLESAGCOEMASVLTGN 800
 Db 736 STSQSTELDSNLSGDPQMRVLCITLQHPGNIIRLWLGRCGLSHECCFDSISLVSSN 795
 QY 801 PHVLVDLTGNALEDGLRLLCOGLRHPVCRILTMLKICRTPTAAACDELASTLSVNQSL 860
 Db 796 QKLVLDLSDNALGDFGIRLLCVGLKHLNCKMLVSCCLTSACQDLASVLSHSL 855
 QY 861 RELDLSNELGLVLLCEGLRHPCKLTOLGLGCRLSAACEGLSVLQANHNIREL 920
 Db 856 TRYVGENALGDSGVAILCEKAKNPQCNLQKGLVNSGLTSVCCSALSSVLTNQNLTHL 915
 QY 921 DLSFNDLGDWGLWLLAEGLOHPACRLQKWLDCGLTAKACENLYFTLGINQTLTDLT 980
 Db 916 YLRGNTLGDGKIGLLCEGLLHPCKLQVLELNCNLTSHCCWDLSTLLTSSQSLKLSIG 975
 QY 981 NNALGDTGVRLCKRLSHPCCKLRVLWLFQMDLNKMTSHSLAALRVTKPYLDI 1033
 Db 976 NNDLGLGVMMFCEVLKQOQSLQNLGLSEMYENYETKSALETQKEKPELTV 1028

RESULT 12

ABB77910

ID ABB77910 standard; protein; 449 AA.

AC ABB77910;

XX 07-OCT-2002 (first entry)

DT Amino acid sequence of human leucine-rich repeat protein HLRBM1.

DE Human; leucine-rich repeat; HLRBM1; proliferative disorder;

DE immune condition; apoptosis; signal transduction; autoimmune disease;

KW haematopoietic cell disease; graft-versus-host disease; allergy; asthma;

KW cardiovascular disorder; neurological disease; pheromone;

KW pulmonary disease; chronic obstructive pulmonary disease;

KW allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;

KW haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;

KW inflammatory disorder; systemic lupus erythematosus;

KW cardiovascular disease; cancer.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Domain 144..159

FT /note= "transmembrane domain"

XX WO200252011-A2.

PN 04-JUL-2002.

PD 20-DEC-2001; 2001WO-US049740.

PP 22-DEC-2000; 2000US-0257773P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder J, Ramanathan C, Mintier G;

XX WPI; 2002-566676/60.

DR N-PSDB; ABL59333.

XX New HLRBM1 nucleic acids for preventing, treating or ameliorating e.g.

PT proliferative disorders, immune conditions, a disorder related to

PT aberrant apoptosis modulation or developmental disorders.

XX Claim 20; Fig 1A-E; 371pp; English.

PS The present sequence represents a human leucine-rich repeat containing

CC protein, designated HLRBM1. HLRBM1 polypeptides and polynucleotides are

CC useful for preventing, treating or ameliorating a medical condition such

CC as a proliferative disorder, immune condition, or a disorder related to

CC aberrant apoptosis modulation, either directly or indirectly, and in

CC modulating signal transduction activity in various cells, tissue and

CC organisms. They are also useful for treating, preventing, or diagnosing

CC diseases of haematopoietic cells, autoimmune disease, graft-versus-host

CC disease, allergic conditions (e.g. asthma), cardiovascular disorders, and

CC neurological diseases, and for increasing the organisms' ability to

CC synthesize and/or release pheromones. The polypeptide may also be used in

CC treating, preventing or ameliorating pulmonary disease (e.g. chronic

CC obstructive pulmonary disease, allergic rhinitis, or bronchial

CC hyperresponsiveness), reproductive disease, haematopoietic disease,

CC platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious

CC disorders (e.g. innate immunity to bacterial pathogens, or adaptive

CC immune response), immune and inflammatory disorders (e.g. systemic lupus

CC erythematosus), cardiovascular diseases and cancers. HLRBM1 nucleic

CC acids may further be used in chromosome identification or mapping, as a

CC chromosome marker, as molecular weight markers, as diagnostic probes, in

CC gene therapy, in raising anti-DNA antibodies, or as antigens for

CC eliciting immune responses

XX Sequence 449 AA;

QY Query Match 42.7%; Score 2334; DB 5; Length 449;

Best Local Similarity 90.7%; Pred. No. 2e-214; 0; Indels 46; Gaps 2;

Matches 449; Conservative 0; Mismatches

QY 228 MNQATSCSMQDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGWCLCW 287

Db 1 MNQATSCSMQDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQGWCLCW 60

QY 288 EKRPTELLNSLRKLLPELSLLITTRPTALEKHLRHLLEHPRHVEILGFSAEKKEYF 347

Db 61 EKRPTELLNSLRKLLPELSLLITTRPTALEKHLRHLLEHPRHVEILGFSAEKKEYF 120

QY 348 KYFHNAGAGQVFNVRDNEPLFTWCFVPLVCWVCTCIQQOLEGGGLLRQTSRTTAV 407

Db 121 KYFHNAGAGQVFNVRDNEPLFTWCFVPLVCWVCTCIQQOLEGGGLLRQTSRTTAV 180

QY 408 YMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLWNQKILPEEQDLRKHGLDGEDVSF 467

Db 181 YMLYLLSLMQPKPGAPRLQPPNQRGLCSLAADGLWNQKILPEEQDLRKHGLDGEDVSF 240

QY 468 LNMNIFQKIDINCERYYSFIHLSFOEFFAAMYIYLDGEAGGAGDQDVTLLTETAYASERS 527

Db 241 LNMNIFQKIDINC-----ERS 255

QY 528 FLATSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAGSDGSLTQOGLSLEFF 587

Db 256 FLATSRFLFGLLNEETRSHLEKSLCWKVSPIHKMDLLQWIOSKAGSDGSLTQOGLSLEFF 315

QY 588 SCLYEIOEEEFIOQALSHPFQVIVVSNIAKMEHWSFCLKRCRSQVHLHYGATYSADG 647

Db 316 SCLYEIOEEEFIOQALSHPFQVIVVSNIAKMEHWSFCLKRCRSQVHLHYGATYSADG 375

QY 648 EDARCSAGAHNTLLVQLRPERVTLDAYSEHLAAALCTNPNIIELSLYNALGSRGVKLL 707

Db 376 EDARCSAGAHNTLLVQLRPERVTLDAYSEHLAAALCTNPNIIELSLYNALGSRGVKLL 434

QY 708 CQGLRHPNCKLQNL 722

Db 435 CQGLRHPNCKLQNL 449

RESULT 13
ADE36452
ID ADE36452 standard; protein; 719 AA.
XX
AC ADE36452;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human PAAD domain associated protein.
XX
KW cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
KW antiatheric; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; pyrin 2; PAAD domain.
XX
OS Homo sapiens.
XX
FN US2003077699-A1.
XX
PD 24-APR-2003.
XX
PF 25-SEP-2001; 2001US-00965621.
XX
PR 26-SEP-2000; 2000US-00671760.
PR 26-SEP-2000; 2000US-0367367P.
XX
PA (REED/) REED J C.
PA (GODZ/) GODZIK A.
PA (CHUZ/) CHU Z.
PA (PAWL/) PAWLOWSKI K.
PA (FIOR/) FIORENTINO L.
PA (ARIZ/) ARIZA M E.
PA (STEH/) STEHLIK C.
XX
PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
PI Stehlik C;
XX
WPI; 2002-471256/50.
DR N-PSDB; ADE36451.
XX
PT Novel isolated PAAD domain containing polypeptide useful for inducing
PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
PT therapy for treating cancer.
XX
PS Disclosure; SEQ ID NO 59; 93pp; English.
XX
CC (I) comprising 80% identity to the amino acid sequence of PAAD and
CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
CC -like protein containing a caspase recruitment domain (ASC)-2 fully
CC defined in specification, where (I) is biologically active. (I) is useful
CC for identifying a (I)-associated polypeptide, an agent altering that
CC association and agents that modulate PAAD domain mediated inhibition of
CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
CC for identifying an agent that modulates the activity of the NB-ARC domain
CC of (I). (I) or its functional fragments is useful in altering cellular or
CC biochemical process such as apoptosis, NFkappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c

CC -Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (I) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
CC therapy. This is the amino acid sequence of human PAAD domain associated
CC protein.
XX
SQ Sequence 719 AA;
Query Match 30.2%; Score 1653.5; DB 5; Length 719;
Best Local Similarity 46.7%; Pred. No. 8.5e-149;
Matches 335; Conservative 125; Mismatches 174; Indels 83; Gaps 10;
QY 11 CELSTYLELENAVELKKKLYLGATTELGEGKIPW----- 45
DB 8 CXLARYLEDELDVLUKTFMHELYPP-QKGCIPLPXXXXXXXKXXXXXXXXXXXXXXX 66
QY 46 -----GSMKAGPLEMAQLLITHFGPEAWRLALS 75
DB 67 XX 119
QY 76 TFERINRXLDERGQEDLVDPQETRYDYRRKFRMLMEDRNARLGEVNLSHRYTRLL 135
DB 120 LLEYLSRISICK-----MKDYRKRYKRVSRFQCTEDRNARLGESVSLNKRYTRL 173
QY 136 VKEHNPQVQOQLDGTGRHARTVGHQASPIKILTEPDEERPEPTVMQGAAG 195
DB 174 IXEHSQEREQELLAIGK--TKTCESVPSPITKMLLELFPDDEHSEPVHTVVFQGAAG 231
QY 196 KSMIAHKVMDWADGKLFQGRFDYLYNCRMNQSAECSSMODLIFSCWPPSPAPLOBL 255
DB 232 KTLARKMMDWASGTLVQDRFDYLYHCRVS-LVTQSRGLDILMCCPDPNPIHKI 290
QY 256 IRVPERLLFIIDGDELKPSFHDQPMWCLCWEKRPTELNLNLSIRKLLPELSLIT 315
DB 291 VRKPSRILFLMDGDELQGFDEHIGPLCTDQKAERGDIILLSSLRKLLPEASLIT 350
QY 316 RPTALEKHLRLEHPRHVEILGFSEAEKRYKYFHNAEQAGOVNYVRDNEPLFTMCF 375
DB 351 RPALEKLOHLLDHPHVEILGFSEAKRYKYFSDQAQAAAFSLIQENEVLFTMCF 410
QY 376 VPLVQWVCTCLOQLGEGGLRQTSRTTAVYMLYLLSLMQPKGAPRLQPPNQRGLC 435
DB 411 IFLVQWVCTGLKQOMESGSLAQTSKTTAVYVFLSLQLPRGSGOEHGLCAHLWGLC 470
QY 436 SLAADGLNNQKILFREQDLRKHGLGDEVSAFLNNNI FOKDINCERYYSFIHLSFOEPPA 495
DB 471 SLAADGIWNNKILFBESDLRNEHGLQKADVSALRNNLFOKEVDCFKFYFIHMTFOEPPA 530
QY 496 AMYVILDEGEGAG-----PDQVTRILATEYAFSERSFALTSRFLGLLNEETS 546
DB 531 AMYLLEREKEGRTNVPGRSLKPLGRDVTLLNKGKFEKGLYIFVVRFLGLVNEETS 590
QY 547 HLEKSLKWKVSPHIRMDLIQWTSKAQSDGSTLQOQSLEFFSCLYEIOEEETIQALSHF 606
DB 591 YLEKLSCKISQIRLELLKLVKAKAKKLOIQSLELFYCLYEMQEDFVQRAWDYF 650
QY 607 QVIVVSNIAKMEHVMVSSFLKRCRKAQVLHUYGATYSADGEDRARCASGAHTLLVQ 663
DB 651 PKIEI-NLSTRMDHVMVSSFCIENCHRVSLSL-GFLHNPKEEKEEGEGRHLDVMQ 705
RESULT 14
AAE21062
ID AAE21062 standard; protein; 344 AA.
XX

AC AAE21062;
 XX 01-JUL-2002 (first entry)
 DT rno (upregulated by nitric oxide)-1 protein.
 DE Upregulated by nitric oxide; NO; rno-1; differentiation; death; cancer;
 XX acute myeloid leukaemia; AML; cytostatic.
 XX Unidentified.
 OS
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 249 /label= Unknown
 FT /note= "This residue is represented in the specification
 FT as O"
 FT Misc-difference 256 /label= Unknown
 FT /note= "This residue is represented in the specification
 FT as O"
 XX US2001029033-A1.
 XX 11-OCT-2001.
 XX 06-MAR-2001; 2001US-00799983.
 XX 06-MAR-2000; 2000US-0186971P.
 XX (SHAM/) SHAMI P J.
 XX (PARK/) PARKER C J.
 XX Shami PJ, Parker CJ;
 XX WPI; 2002-009982/01.
 XX Novel polypeptide regulated by nitric oxide useful for inducing acute
 XX myeloid leukemia cell differentiation and apoptosis.
 XX Claim 29; Fig 7B; 23pp; English.
 XX The present invention relates to novel genes which are upregulated by
 XX nitric oxide (NO), designated as rno and their corresponding proteins.
 XX The invention also relates to the isolation and characterisation of three
 XX isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
 XX are useful for inducing differentiation and death in a cancer cell.
 XX Diagnosing acute myeloid leukaemia (AML) comprises detecting the
 XX expression of a rno gene comprising the nucleotides coding for rno
 XX protein. The present sequence is rno-1 protein
 XX Sequence 344 AA;
 SQ
 Query Match 30.2%; Score 1652; DB 5; Length 344;
 Best Local Similarity 96.9%; Pred. No. 3.6e-149;
 Matches 316; Conservative 1; Mismatches 7; Indels 2; Gaps 1;
 QY 710 GLRHPCNQLRKRKCRSSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 769
 DB 21 GLQOP--RRRLWKRCRISACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLR 78
 QY 770 HPQCRQLQMIQRLKQLESACQEMASVLTGPHVLDLTGNALDGLRLCCQGLRHPV 829
 DB 79 HPQCRQLQMIQRLKQLESACQEMASVLTGPHVLDLTGNALDGLRLCCQGLRHPV 138
 QY 830 CRURTWLKTCRTAAACDLASTLSVNSQSRRLDLSNLGLDGLVLLCEGLRHPTCKL 889
 DB 139 CRRLTLWLKTCRTAAACDLASTLSVNSQSRRLDLSNLGLDGLVLLCEGLRHPTCKL 198
 QY 890 QTLRLGTCRGSACGLSVLQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKL 949
 DB 199 QTLRLGTCRGSACGLSVLQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKL 258
 QY 950 WLDSCGLTAKACENLYFTLGINTLTDLYLTNNALGDTGVRLICKRLSHPGCKLRVLWLF 1009

DB 259 WLDSCGLTAKACENLYFTLGINTLTDLYLTNNALGDTGVRLICKRLSHPGCKLRVLWLF 318
 QY 1010 GMDLNKMTSHRLAALRVTPYLDIGC 1035
 DB 319 GMDLNKMTSHRLAALRVTPYLDIGC 344
 RESULT 15
 AAO15593
 ID AAO15593 standard; protein; 994 AA.
 XX AC AAO15593;
 XX 31-OCT-2002 (first entry)
 XX Human PYRIN-3 protein.
 DE Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-3.
 XX Homo sapiens.
 XX WO200261049-A2.
 XX 08-AUG-2002.
 XX 31-JAN-2002; 2002WO-US002967.
 XX 31-JAN-2001; 2001US-0265231P.
 XX 10-SEP-2001; 2001US-0318645P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX (AMHP) WYETH.
 XX Bertin J, Wang W, Blatcher M;
 XX WPI; 2002-627477/67.
 XX N-PSDB; AAL44366.
 XX New PYRIN polypeptides and nucleic acids useful for modulating and
 XX diagnosing stress-related, apoptotic and inflammatory responses, or for
 XX treating inflammatory and immune system disorders, cancers, or
 XX neurological diseases.
 XX Claim 8; Fig 11; 167pp; English.
 XX The invention comprises the amino acid and coding sequences of human
 XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 XX useful for modulating and diagnosing stress-related, apoptotic and
 XX inflammatory responses. The PYRIN protein and DNA sequences are useful
 XX for treating: inflammatory disorders and immune system disorders (e.g.
 XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
 XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
 XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
 XX (e.g. systemic lupus erythematosus and arthritis); and neurological
 XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 XX protein and DNA sequences may also be used in screening assays, detection
 XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 XX predictive medicine (e.g. diagnostic assays, clinical trials and
 XX pharmacogenomics) and transcription profiling. The present amino acid
 XX sequence represents the human PYRIN-3 protein
 XX Sequence 994 AA;
 SQ
 Query Match 26.6%; Score 1457.5; DB 5; Length 994;
 Best Local Similarity 33.9%; Pred. No. 9.2e-130;
 Matches 353; Conservative 197; Mismatches 402; Indels 89; Gaps 19;

Search completed: July 30, 2004, 13:49:22
Job time : 61 secs

QY 13 LSTYLELEAVELEKFKLYLGAT- ELGEGKI PWGMEKAGLEMAQLLITHFGERAWR 71
Db 11 LMWYLEELKEEFKFKHEHLKQMTLQELKQIPWTEVAKSREELANLLIKHYEQQAWN 70
QY 72 LALSTFERINRKOLWEGQREDLVRDPOETRYDYRRKF-RLMEDRNARLGCYNLSHRY 130
Db 71 ITLAIQKMDRKDLCKVMRER--TGYTKTYQAHAKQKFSLSWSSKV----- 116
QY 131 TRLLLVKHSNPMOVQOQLLDTGRGHARTVGHQASPIKIETLFPDDEPPEPPTVMQ 190
Db 117 TEIHLFYEE---EVKQEC-----HLDRLEAP-KETGKQPRTVIIQ 155
QY 191 AAGIGKSLAHKVMLDWADGKLFGQRPDYLYINCREMNQSAECSDMLFSCWBPESA 250
Db 156 POGIGKTTLLMKLWMSNDKIFDRFLYTFYFCCREURE-LPPTSGLADLISREWPOFAA 214
QY 251 PLOBLIRVPERLLFTIDFDELKPSFHDPOGPWCMLCWEKRPTELLNLSLRKLLPELS 310
Db 215 PITEIVSQPERLLFVIDSFEELQGLNEPDSDLCDLMEKRPVQVLLSSLRKMLPEAS 274
QY 311 LLITRPTALEKHLRHLHLEPHRVEILGFSEARKEYFYKYPHNAEQAGQVYVRDNEPL 370
Db 275 LLIAIKVPCKELRDQVTISIIYQPRGFNEDRLVYFCCFEKPKRAMEAFNLVRESOL 334
QY 371 PTMCFVPLVCHVCTCLOOQLEGGGLRQTSRTTAVMYLLSLMOPKPGA--PRLOPP 428
Db 335 FSIQOILPILLCWILCTSKQEMQKQKDALTCOSTTSVTSFVFNLFPE-GAEGFTPTQ 393
QY 429 PNQRLGLSLAADGLWNQKILFEEQDLKXHLGDGEDVSAFLNMNIFQKINCERYYSFIHL 488
Db 394 HOLKALCSLAEGNWTDTFECEDDLRENGVVDADIPALLGTILLKYGERESSVFLHV 453
QY 489 SPQEFFAAMYI-----LDEGGGAGPDQDVTRLLTEVAFSERSFALTSRFLGLNEET 544
Db 454 CIOEFCAALFYLLKSHLDHPHAPVRCVOEL--LVANFEKARRAHWIFLGCELTGLLNKKE 511
QY 545 RSHLEKSICWKVSPHIMWDLLOWKQASQDSTLQOGLFEFFSCLYEIOEEFTQOALS 604
Db 512 QEKLDADFGLQSEIQQIHHQCLKSGERNGPQGVDSLAIFYCLFEMQDPAPVKQAVN 571
QY 605 HFQVI--VVSNIASKMEHMYSSFCILKRCRAQVILHYGATYSADGEDRARCAGHTLL 661
Db 572 LLOEAFNHIIIDNV---DLVVSAYCLKYCSSURKLCF--SVQNVFKKEDHSSTSYSLI 625
QY 662 VOLRPERTVLDAYSEHUALAALCTNPNIELSLYRNALSGRGVLLCOGLRHPNCKLQNL 721
Db 626 CW-----HHICSVLTTSGLHRELQVQDSTLSESTFVTCNQLRHPSCRLQKL 672
QY 722 RLKRCRISSAC-----EDLSAALIANKLTRMDLSGNGVGFPGMMLLCGLRHPQ 772
Db 673 GINNVSFGQVLLFEVLYQPDULKYLSFTYTKLSRDDIRS-----LCDALNTPA 722
QY 773 CRLQMIQLRKQOLESAGQEMASVILGNPHILVELDTGNALEDLGLRLLCQGLRHPVCR 832
Db 723 GNVKELALVNCHLSPIDCEVLGLLITNNKLTLYNVSCNL-DTGVPILCEALCSPDTVL 781
QY 833 RTLWIKICRLTAACDELAFTLSVNOISRELDSLNEIGDLGVLLCEGLRHPCKLOT 892
Db 782 VYLMALFCHLSEQCCEYISEMLLRNKSRYLDLSANVLKDEGLKTLCEALKHPDCCDLSL 841
QY 893 RLIGICRLGSAACEGLSVVLOANHNRLDLSFNDLDGMDGLMLAELGLOHPACRLQKWL 952
Db 842 CLVKCFYTAACEDIASALISQNLKILOIGCNEIGDVGVQVLLCRLALHTDCLREIIGLE 901
QY 953 SCGLTAKACENLYFTLGINQTLTDLVLTNNALGDTGVRLLCKRISHPGCKRLVWLFGMD 1012
Db 902 ECGLTSTCKDLASVLTCSTKITLQNLNLTLNLTLDHTGVVVLCEALRHPICALQVGLRKTD 961
QY 1013 LNKWTHSRALARVTKPYLDI 1033
Db 962 FDEETQALLTAEERPNPLTI 982

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:45:21 ; Search time 22 Seconds

(without alignments)

4525.374 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472

Sequence: 1 MLRTAGRDGLCLSTYLEEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314.5	24.0	1192	2 T17255	hypothetical prote
2	1160.5	21.2	1111	2 A59000	mater protein [imp
3	724	13.2	461	2 A31858	ribonuclease-angio
4	717	13.1	456	2 S20597	ribonuclease inhib
5	709	13.0	456	2 A31857	ribonuclease inhib
6	494.5	9.0	483	2 S27880	Nasopressin recept
7	377.5	6.9	1130	2 A48843	MHC class II trans
8	258	4.7	1004	2 T31665	hypothetical prote
9	216.5	4.0	1075	2 T31668	hypothetical prote
10	183.5	3.4	1232	2 A55478	neuronal apoptosis
11	180	3.3	312	2 B97746	hypothetical prote
12	179.5	3.3	545	2 T52068	RAN GTPase-activat
13	178	3.3	533	2 T52063	ran GTPase-activat
14	177.5	3.2	506	2 A45841	T-complex-associat
15	172.5	3.2	618	2 T48193	hypothetical prote
16	171	3.1	1447	2 T42628	hypothetical prote
17	170	3.1	568	2 F86291	hypothetical prote
18	169.5	3.1	789	2 T52067	hypothetical prote
19	169.5	3.1	1121	2 T02764	myosin-I binding p
20	162	3.0	631	2 C89243	protein F28C1.3 [i
21	162	3.0	631	2 T21471	hypothetical prote
22	161.5	3.0	589	2 A36983	RNAL homolog fugi
23	160	2.9	998	2 T23427	hypothetical prote
24	158	2.9	535	2 T48102	RAN GTPase activat
25	156.5	2.9	589	2 T52070	RNAL protein homol
26	155	2.8	587	2 JC5300	Ran GTPase activat
27	152.5	2.8	2493	2 A55481	adenylate cyclase
28	152	2.8	1010	2 T36383	probable large ATP
29	142.5	2.6	1253	2 T45787	disease resistance

30 142 2.6 934 2 T05201 hypothetical prote
31 140.5 2.6 1389 2 T13852 gene wheeler prote
32 140 2.6 1039 2 A85096 hypothetical prote
33 139.5 2.5 271 2 D84586 hypothetical prote
34 138 2.5 526 2 C84552 hypothetical prote
35 137.5 2.5 1237 2 A81915 hypothetical prote
36 137.5 2.5 1385 2 T13887 tlr protein - frui
37 136.5 2.5 1286 2 T33476 hypothetical prote
38 135.5 2.5 1066 2 T15864 hypothetical prote
39 135 2.5 592 2 T52139 LRR-containing F-b
40 135 2.5 1217 2 T52348 disease resistance
41 134.5 2.5 906 2 G96621 probable disease r
42 134.5 2.5 957 2 E84547 probable disease r
43 134 2.4 607 2 E96598 protein F20N2.2 [i
44 133 2.4 1110 2 F84547 probable disease r
45 132.5 2.4 907 2 J80176 orphan G protein-c

ALIGNMENTS

RESULT 1

T17255

hypothetical protein DKFZp586O1822.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T17255

R:Koehler, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18722

A:Accession: T17255

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-1192 <ROE>

A:Cross-references: EMBL:AL117470

A:Experimental source: adult uterus; clone DKFZp586O1822

C:Genetics:

A:Note: DKFZp586O1822.1

Query Match 24.0%; Score 1314.5; DB 2; Length 1192;
Best Local Similarity 39.9%; Pred. No. 4.6e-94;
Matches 323; Conservative 119; Mismatches 293; Indels 75; Gaps 19;

QY 133 LLLVKEH--SNPMQVQO--QLLDTGRGHARTVGHQASPIKIETLFEDEEPEPEPRTVV 187
DB 1 LLLQRPSPQDPLVKRSWPDYVEENRGL-----IEIRDLFGPLDQEQ-PRIVI 50
QY 188 MQGAAGIGKSLAHKVMLDWADGKLFQGRFDYLFYINCRMNQOSATECSMODLIFSCWEE 247
DB 51 LQGAAGIGKSLARQVKEAWGQGLYGRFQHVFFSCRELAQSKV-VSLAELIGKDGTA 109
QY 248 PSAPLOELIRVPERLLFIIDGDELKPSFHPDQGPWCWCEKRPTELLNSLIRKLLP 307
DB 110 TPAPIRQILSRPERLLFIIDGDELKPSFHPDQGPWCWCEKRPTELLNSLIRKLLP 169
QY 308 ELSLLITTPALEKHLLEHPHVEILGRSEAEKKEYFYKFFHNAQAGQVFNVRDN 367
DB 170 EASFUIITARTALQNLIPSEQARWVEVLGFSSSRKEYFYRYFTDQAIRAFRLVKN 229
QY 368 EPLFTMCFVPLVCWVCTCLOQLEGGLLRQTSRTTTAVVYMLLSLMQKPGAPRLQP 427
DB 230 KELWALCLVFPVSWLACTCLMQMKRKEKLTLSKTTTTLCHYLAQALQAPLPQL-- 287
QY 428 PPNQRLGSLAAGLWNNOKILFEEDLRKHGDLGDSVAFNLNMIFQKDCINERYYSFIH 487
DB 288 ----RDLCSLAAGIWKQKTLFSPDDLKRGHLDGAIISTFLKMGILQEH-PIPLSVSFIH 342
QY 488 LSPQEFFFAAMYILDGEGGAGPDQ----DVTRLITTEYAFSERSPLATSTRFLGLNNEE 543
DB 343 LCFQEFFFAAMSVYL--EDEKGRGKHSNCITIDLEKTLLEYAGI-HGLFGASTTRFLGLLSD 400
QY 544 TRSHLEKSLCWKSPHMKMDLLOWIQSKAQSDGSTLQQQSLEFFFCFLYEIOEEETIQAL 603

Arch. Biochem. Biophys. 312, 421-428, 1994
A:Title: Purification and characterization of human brain ribonuclease inhibitor.
A:Reference number: S48636; MUID:94311593; PMID:8037455
A:Accession: S48636
A:Status: preliminary
A:Molecule type: protein
A:Residues: 2-14 <NAD>
R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.
submitted to the Protein Sequence Database, March 2000
A:Reference number: 224374
A:Accession: T47188
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <AAA>
A:Cross-references: EMBL:AL161967
A:Experimental source: adult testis; clone DKFZp434K249
C:Genetics:
A:Gene: GDB:RNH
A:Cross-references: GDB:125274; OMIM:173320
A:Map position: 11p15.5-11p15.5
A:Note: DKFZp434K249.1
C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C:Keywords: blocked amino end; duplication
F:34-433/Region: leucine-rich 57-residue repeats
F:57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>
F:114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>
F:174-195/Region: inhibitory
F:228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>
F:285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>
F:399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Query Match 13.2%; Score 724; DB 2; Length 461;
Best Local Similarity 42.5%; Pred. No. 1.7e-48;
Matches 150; Conservative 64; Mismatches 139; Indels 0; Gaps 0;

Qy 671 LLDAYSEHLAALCTNPNLIELSYRNALSGRGVLLCOGLRHPNCKLQNLRLKRCRISS 730
Db 40 LTEARCKDIISALRNPALAEINLSNELDGVGVHCVLQGLTPSCIKQLSLQNCCLTG 99
Qy 731 SACEDLISALANKNLTRMDLSGNGVGPFGMMLCEGLRHPQCRLOMIQLRKCOLESAC 790
Db 100 AGCGVLSLTPLTQELHLSNLLDAGLQGLCEGLDPCRLKLEKLEKLEKLEKLEKLEK 159
Qy 791 QEMASVLGTNPVHVELDTGNALDGLVLLCEGLRHPVCRRLTLWLKICRLTAACDEL 850
Db 160 EPLASVLRAKPDFKELTVSNNDINEAGRVLCQGLKSPQLEALKLESQGVTSNCRDL 219
Qy 851 ASTLSVNQSLRELDLSNLELDTGNALDGLVLLCEGLRHPQCRLOMIQLRKCOLESAC 910
Db 220 CGIVASKASLRELALSGNKGIDVGNALCPGLLHFSRLTLWIWECGITAKGCGDLCEV 279
Qy 911 LQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKWLDSGLTAKACENLYFTLGI 970
Db 280 LRAKESLKEKSLAGNELDGLRHPVCRRLTLWLKICRLTAACDEL 339
Qy 971 NOTLTDLYTNALGDTGVRLLCRLSHPGCKLRVLMFGMDLNKMTWHSRLAA 1023
Db 340 NFFLELQISNNRLDAGVRELQCGIGQSGVLRVLMWADCDVSDSSCSLAA 392

RESULT 4
S20597
Ribonuclease inhibitor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
C:Accession: S20597
R:Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
Biochim. Biophys. Acta 1129, 335-338, 1992
A:Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
A:Reference number: S20597; MUID:92162755; PMID:1536887
A:Accession: S20597
A:Molecule type: mRNA
A:Residues: 1-456 <AA>

A:Cross-references: EMBL:X62528; NID:957670; PIDN:CAA44388.1; PID:957671
C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.1%; Score 717; DB 2; Length 456;
Best Local Similarity 41.5%; Pred. No. 5.9e-48;
Matches 147; Conservative
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-456 <HOF>
R:Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
Biochemistry 29, 8827-8834, 1990
A:Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclease
A:Reference number: A35830; MUID:91104783; PMID:2271559
A:Accession: A35830
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 82-456 <VIC>
A:Cross-references: GB:M8700; NID:9164638; PIDN:AAA63448.1; PID:9164639; GB:J02925
C:Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
C:Keywords: liver
F:280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 13.0%; Score 709; DB 2; Length 456;
Best Local Similarity 42.7%; Pred. No. 2.5e-47;
Matches 151; Conservative 60; Mismatches 143; Indels 0; Gaps 0;

Qy 671 LLDAYSEHLAALCTNPNLIELSYRNALSGRGVLLCOGLRHPNCKLQNLRLKRCRISS 730
Db 35 LTEEHCKDIGSALRANPSLTCLRTNELDGLDAGVHLVQLQGLSPCTCKIQLSLQNCCLTE 94
Qy 731 SACEDLISALANKNLTRMDLSGNGVGPFGMMLCEGLRHPQCRLOMIQLRKCOLESAC 790
Db 95 AGCGVLSLTPLTQELHLSNLLDAGLQGLCEGLDPCRLKLEKLEKLEKLEKLEKLEK 154
Qy 971 QEMASVLGTNPVHVELDTGNALDGLVLLCEGLRHPVCRRLTLWLKICRLTAACDEL 850
Db 335 NSSLFELQMSNPLDGLSGVVELCKALGYDPTVLRVLMWADCDVSDSSCSLATV 388

RESULT 5
A31857
Ribonuclease inhibitor, hepatic - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
C:Accession: A31857; A35830
R:Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
Biochemistry 27, 8537-8544, 1988
A:Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals the
A:Reference number: A31857; MUID:89118268; PMID:3219361
A:Accession: A31857
A:Molecule type: protein
A:Residues: 1-456 <HOF>

Db 155 EPLASVLRATRALKELTUNNDIGEARVLGQGLADSAQOLETTLRENCGLTPPANCODL 214

QY 851 ASTLSVNSQISRELDLSNELDGLVLLCEGLRHPCKQLTLRLGICRLGSAACEGLSVV 910

Db 215 CGIVASQASRELDLSNGSLGAGIAELCPGLSPASRLKTLWLWECDDITASGCDLRCV 274

QY 911 LOAHNHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKWLDSOGLTAKACENLYFTLGI 970

Db 275 LQAKETLKSLAGNKLGBEGARLLCESLLOQCQLESJWVKSCTITACQCHVSLMLTQ 334

QY 971 NOTLTDLYTNALGDTGVRLLCRLKLSHPCKLRLVWLFGMDLNKMTHTSRLAAL 1024

Db 335 NKHLLEQLSSNKLGDGSIQELCOALSQPGTTLRLVLCLGDCVETNSGCSLASL 388

RESULT 6

S27880

Nasopressin receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 05-Nov-1999

C:Accession: S27880

R:Herrera, V.L.; Ruiz-Opazo, N.

submitted to the EMBL Data Library, February 1992

A:Description: Characterization of a cDNA encoding an AII and AVP receptor isoform.

A:Reference number: S27880

A:Accession: S27880

A:Molecule type: mRNA

A:Residues: 1-483 <HER>

A:Cross-references: EMBL:M85183; NID:g202805; PIDN:AAA03623.1; PID:g202806

C:Genetics:

A:Gene: AVP

Query Match 9.0%; Score 494.5; DB 2; Length 483;

Best Local Similarity 28.8%; Pred. No. 1.6e-30; Indels 85; Gaps 10;

Matches 147; Conservative 74; Mismatches 205;

QY 391 LEGGGLRSTSTTTAVMYLLSMQPK-PCAPRLQPPNPORGSLAAGLWNGKILF 449

Db 1 MELGRDLSRTSTTISVYLLFITSMLKSAGTNGPRVQ--GELRMLCRLAREGLKHQAQF 58

QY 450 EQODLR--KHGLDGBDV-SAFINMFIQKINCERYYSFIHSPQFFFAAMYIIDELEG 506

Db 59 SEKDLERLKLGQSQVQTMFLSKKELPGVLETVVVYQIDQSOFORFLAALSYYLLD-ABG 117

QY 507 GAGPDQDVTLLTEYAFSERSFLATSLRFLGLNNEETSHLEKSLCWKVSPIHKMDLQ 566

Db 118 AFNGSAGSVQMLNDSAGLRGHALTTRFLFGLLSTERTRDIGNHFGCVVPRGVQDITLR 177

QY 567 WTQSKAQSDGSTL-----QQGSLEFFSCLYRIQESEFIQ 601

Db 178 WYQGQSQPKVATVGAEEKKDELKDEABEEEEEEELNFGLELLYCLYETQEDDFVQ 237

QY 602 ALSHFQVIVVSN-ASKMEHWYSSFCCLKCRSAQVILHYGATYSADGEDRAR----- 652

Db 238 ALSSLPFWLERVTRMDLEVLSCVQCPCQALRLVSCGLVAAKEKKKKKSGFMNRL 297

QY 653 ---CSAGAHLLVQLRPVTLVDAYSEHLAAALCTNPNLIELSLYRNALSGRGVKLLCQ 709

Db 298 KGSQSTGQPPASILRP-----LCE 317

QY 710 GLRHNCKLQNLRLKRCISSACEDLSAALIANKNLRTMDLSGNGVGFPGMMLLCEGLR 769

Db 318 AMITQQCGLSILTLSHCKLPDVCARDLSEALKVAPSLRELGLQNLTEAGRLLSQGLA 377

QY 770 HPQCRLQMLQKCKOLESACQEMASVLGTNPHVLELDLTGNALEDGLRLLCQGLRHVP 829

Db 378 WPKCKVQTLRIQMPGLQE-VIHYLVIVQQSPVLTTDLSCQLPGTVPVEPLCSALKHPK 436

QY 830 CLRRTLWLKICRLTAAACDELASTLSVNQSL 860

Db 437 CGLKTLSLTSVELTENPLRELQAVKTLKPD 467

RESULT 7

A48843

MHC Class II transactivator - human

N:Alternate names: CIITA

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 08-Oct-1999

C:Accession: A48843

R:Steimie, V.; Otten, L.A.; Zufferey, M.; Mach, B.

Cell 75, 135-146, 1993

A:Title: Complement activation cloning of an MHC class II transactivator mutated in hereditary

A:Reference number: A48843; MUID:94006536; PMID:8402893

A:Accession: A48843

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1130 <STE>

A:Cross-references: GB:X74301; NID:9414112; PIDN:CAA52354.1; PID:9414113

Query Match 6.9%; Score 377.5; DB 2; Length 1130;

Best Local Similarity 23.6%; Pred. No. 8e-21;

Matches 220; Conservative 122; Mismatches 355; Indels 235; Gaps 40;

QY 92 EDLVRDPQETY-----RDYVRKFLMEDRNARLGEVCNLSHRYTRLLLVKHSNMQV 145

Db 339 EQFYRLQDITYGAEPAGPDGILVEVDLVQ---ARL-----ERSSSKSLERELATPDWA 388

QY 146 CQQLDVTGRHARTVGHQASPIKIEITLPEDEERPEPPRTVVMQGAAGICKSLAHKVM 205

Db 389 ERQLAQGGIAEV-----LLAAKEHRRPRETRVIAVLGKAGQKSTWAGAVSR 435

QY 206 DWADGKLFQGRFDYLYINCREMNQSATSCMQDLIFSCWPPSPAPLQE---LIRVPER 261

Db 436 AMACGRLPQ--YDFVFSVPCNLNRPDAYGLQDLLFSLGQPLVAADVFVSHILKRPDR 493

QY 262 LLIIDGDELKPS---FHDPOGPMCLCWEEKRPTEL--LLNSLRKLLPELSLITTR 316

Db 494 VLLIIDAFAEEAQAQDGLFHSCTGP-----APAEPCSLRGLLAGLFOKLLRGCTLLTAR 548

QY 317 P-----TALEKLRLHLEHPRHVEILGFSAEERKEYFYKFHNA---EQAGQVNVYVDNE 368

Db 549 PGRILVQSLSKADALF-----ELSGFSMEQAQVVMRYFESSGTEHODRALTLURDP 602

QY 369 PLFTMCFVPLVCMVCTLCQOQLEGGLRQTSRTTAVMYLLYLSLMQPKFAPRLQPP 428

Db 603 LLLSHSHPTLCRAVCQLSEALLEGEDAKLPS-TLTGLYVGLL-----GRAALDSP 653

QY 429 PNORG-LCSLAAD-GLWQKILPEEQ-----DLR-----KHGLDGEDVSFAFNANIPOKDI 478

Db 654 PGALAEALAKLAWELGRHQSTLQEDQFPFSADVRTWAMAKGLVQHPRA----- 701

QY 479 CERYYSFTHLSFQEFFAAMYIILDEGEGAGPDQDVTRLLTEYAFSER---SFLALTSRF 535

Db 702 AESELAFFSFLQCFGLALAL-----SGEYKDKELPQYALTPKRPYDNWLEGVPRF 757

QY 536 LFLGLNEETRSHLEKSLCWKVSPIHKMDLLOWIQS---KAQSDGSTLQOQSGLEFFSCLYE 592

Db 758 LAGLIFQPPARCLGALL--GPSAAASVDRKQKVLARYLKRLOPGLTRARQLLELLHCAHE 815

QY 593 IQEESFIIQALSHFQVIVVSNIAKMEHWVSSFCCLKRCESAQVILHYGATYSADGEDRAR 652

Db 816 ABEAGIWHQ-----VVQELPGR-----SFLGTRLTTPDA-HVLGKALEAAGOD--- 858

QY 653 CSAGAHLLVQLRPVTLVDAYSEHLAAALCTN--PNLIELSL---YRNALSGRGVKLL 707

Db 859 -----FSLDLR-----STGICPSGLSGILVGLSVCTVTRFAALSD--TVAL 895

QY 708 CGGLR-HPNCKL-----QML-RLKRCRISSSACEDLSAALIANK 744

Db 896 WESLRHQETKLLQAAEBKFTIEPPFAKSLKDVDELGLVQVQTRRSSSEDTAGELPAVR 955

QY 745 NLTMRDLSGNGVGFPGMMLLCEGLRHPCQLQMIQLRKQLESACQEMASVLGTNPHLV 804

Db 956 DLKLEFALGPVSGQAF-----PKLV 977

QY	805	ELDTGNALEDIGLRLLCQGRHPVCELRITLWLKICELTAACACDELASTLSVNCQSLRELD	864
Db	978	RIITAFSSQLHDLDALEN-----KIGDEGVQSISATFPQKSLUETLN	1021
QY	865	LSLNEIGDLGVLLICEGRHPTCKIQTLRLG-----ICRLGSAACEGLSVVLOAHNHLRE	919
Db	1022	LSQNNITDLGAYKLAEL--PSLAASLLRLSYNNCLCDVGA---ESLARVLPMVSLRV	1076
QY	920	LDLSFNDLGDGWLWLLAELGLQH--PACRLQKLIW	950
Db	1077	MDVQYNKFTAAGAQQLAASLRRCPEHVETLAWW	1108
RESULT 8			
T31665			
hypothetical protein COSI.4 - sea squirt (Ciona intestinalis)			
C/Species: Ciona intestinalis			
C/Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 20-Jun-2000			
C/Accession: T31665			
C/Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lemnard, N.; Tweedie, S.			
submitted to the EMBL Data Library, October 1996			
A:Reference number: Z21050			
A:Accession: T31665			
A>Status: preliminary; translated from GB/EMBL/DBJ			
A:Molecule type: DNA			
A:Residues: 1-1004 <BIR>			
A:Cross-references: EMBL:Z80904; PIDN:CA802586.1			
C:Genetics:			
A:Introns: 30/1; 539/3; 593/3; 662/3; 728/3; 769/3; 817/2; 881/2; 960/1			
Query Match 4.7%; Score 258; DB 2; Length 1004;			
Best Local Similarity 20.3%; Pred. No. 1.5e-11;			
Matches 218; Conservative 157; Mismatches 361; Indels 340; Gaps 52;			
QY	150	LDTCGRHARTVGHQ---ASPIKIETLEPDE-----	181
Db	4	VEFGSAHAAVKIHQHPPEPAPVNPVLVTDQLFKALENAOCYTEDQSAEYAKVYIER--	61
QY	182	PPRTVMQGAAGIGKS-----MLAHKVMLWDADKLFQGRFDYLFYINCEMNQSA	232
Db	62	HANTVVVVGPHGGKTKLLKMMVQILKHELLPD-----TEYFFIYAKDIDFNK	111
QY	233	TECSMODLI-----FSCWPEPSAPLOELIRVPERLLFIIDGFDLKPFSHDPOGPWCIC	286
Db	112	EMTLLEFLTTSNRVKVNYTEESKALITFLHNNPNVAIFFDGLDE--ASLXELVGGYSIC	169
QY	287	W--BEKRPTTELLNSLRKLLPELSLITTPRTALEKLHLLHPRHVEILGSEAEK	344
Db	170	KLDEKSKPEVDIMKN--LFLNALLPRAKIVVTSTPDEMFLNLCYRPTSTFEVLGFLUEAKN	228
QY	345	EYFYKYFHNAOAGOVFNVR---DNEP--LFTMCFPVPLVCCVTCLOOQEGGGLLRQT	400
Db	229	NL-----GTQLGGEKYPAIKILLDQFNLAHLCYLPINFILIVFCLSN--EGSDI----	277
QY	401	SRTTTAVMYLLSLMQ-----PKPGAPRLQPPNORGLCSLAADGLWNOKILPE	450
Db	278	-KTMTQVLIFSMTRFVESHSLKGEVPLDKVGAEMVK-----LACLAYKGLQKRLVFE	329
QY	451	EQDLRKXGLDGE-----DVSFAFLNMIFQKIDINCERYYSFIHLSFQEFFRAAYIY	500
Db	330	KTFDFDVKLADMTVNTFFHYVDISGIRIKILEGN---KRSY--FTTHIWOEFYAAVYIM	385
QY	501	LDEBGGGAGPDQDVTRLTTEYAFSRSFLATLSRFLGLLN-----	541
Db	386	LFVSY-----REFQLKTIIFEDTQSVVV---KMFGLICNPWAYKQLKIIFPATMIKY	436
QY	542	ETBSRSHLEKSLCWKYSPIKMDLLQ---WIOSKAQSDGSTLOQGSLEFPSCIL-----	590
Db	437	BEKKKPFLESWMWESLSAKSEDLIRRFGLWHE--YNDDET---SKKKFDCLPVGLKMGV	490
QY	591	-YEIQEEEFIOALSHF-----QVI-----VVSNTAS	616
Db	491	PKHLPEVDLVLVAKLSFTKPHKRLRLSNWTTTTEVLETILRGHGTITITRFTVNIEM	550

[illegible]

Db 221 QMYKLPDYPRTSIPFVLGLLEBAKNNL-----GTLOGEKYPAIKKILDDQNPNAHLIC 274
 QY 375 FVPLVCVVCTCLOQLEGGGLLRQTSRTTAYMYLILSLMQ-----PKGAPR 424
 Db 275 YLPIFNILFVCLSN-EGSDI-----KTMTQVLIFFSMTRFVLSHLKGEVPLDKVGAEM 328
 QY 425 LOPPPNQGLCSLAADGLWNOKILFEEQDLRKHGLDGE-----DVSFAFLNNIFQ 474
 Db 329 VK-----LARLAYKGLQORUKLVFEKTFDDVKLADWMTNFFHYTVDKSSGIRMKILE 381
 QY 475 KOINCERYYSFIHLSQFQFAAMYIYLDGEGGAGPDQDVRTLLTYAFSERSFLAL--- 531
 Db 382 GN---KRSY-FTHLIWQEFVAAYVLM-----FVSYPEFEQLKPI 417
 QY 532 -----TSRFLGFLNNESTRHLEKSLCWKVP-----HIKMDLLOQWIOSKAQSDG 576
 Db 418 PKDAQWKRVRVGMFGICNPAYKOLK-----LVFPATMIKDYBEKKELM--VPMMEESLW 470
 QY 577 STLOQGSLEFFSCLYEIOBEETQOALSHFQV---IWSNIAASKMEHWSFCLKCRSA 633
 Db 471 SARGEDLIRFGWLHNEYNDDESKFEDYLPVGLKMDAPKHLSEVKDLV--YALKSFTKP 528
 QY 634 QVHL--YGATYSADGEDRARCAGAHLLV-----QLRP 666
 Db 529 HKRLDSYETTTEVLETLRGVHTTTITRFINNIEMKOSLMELLHLHLDAMBEELR 588
 QY 667 ERTVLDAYSEHLAALCTNPNIELSLYNAL--GSGVKLL--COGLRHPNCKLQNLRL 723
 Db 589 IRVTNLSPMYERLSNAINORSNKIQVLVKKHLHDDYDKYLAGCLG-----NISLLYM 642
 QY 724 KRCRISSACEDLSAAL 740
 Db 643 WGTDISSDQCSVLKQAI 659
 RESULT 10
 A55478
 neuronal apoptosis inhibitory protein - human
 N:Alternate names: NAIP
 C:Species: Homo sapiens (man)
 C>Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 02-Feb-2001
 C:Accession: A55478
 R:Roy, N.; Mahadevan, M.S.; McLean, M.; Shutler, G.; Yaraghi, Z.; Farahani, R.; Baird, S.
 d, T.O.; de Jong, P.J.; Surh, L.; Ikeda, J.E.; Korneluk, R.G.; Mackenzie, A.
 Cell 80, 167-178, 1995
 A>Title: The gene for neuronal apoptosis inhibitory protein is partially deleted in indi
 A:Reference number: A55478; MUID:95112344; PMID:7813013
 A:Accession: A55478
 A:Molecule type: mRNA
 A:Residues: 1-1232 <ROY>
 A:Cross-references: GB:U19251
 C:Genetics:
 A:Gene: GDB:SNAG; SMA
 A:Cross-references: GDB:120378; OMIM:600354; OMIM:253300
 A:Map position: 5q12.2-5q13
 C:Keywords: apoptosis; ATP; glycoprotein; nucleotide binding; P-loop; transmembrane prob
 F:94-110/Domain: transmembrane #status predicted <TM1>
 F:470-477/Region: nucleotide-binding motif A (P-loop)
 F:479-496/Domain: transmembrane #status predicted <TM2>
 F:476/Binding site: ATP (Lys) #status predicted
 F:618,823,923,1035/Binding site: carbohydrate (Asn) (covalent) #status predicted
 Query Match 3.4%; Score 183.5; DB 2; Length 1232;
 Best Local Similarity 21.3%; Pred. No. 1.3e-05;
 Matches 148; Conservative 94; Mismatches 248; Indels 205; Gaps 35;
 QY 188 MQGAGIGKSMIAHVKMLDWDGKL--FQGRFDYLFYINCR-----EMN 229
 Db 468 VEGEAGSGKTVLLKKIAFLWASGCCPLNRFQVFLSLSTRPDEGLASIIICDQLEKE 527
 QY 230 QSATCSMODLFSWCPSPAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLWSE 289
 Db 528 GSVTECMENII-----QQL---KNQVFLDLDYKEI---CSIPQ----- 561

QY 290 KRPTTELLNSLRKKLLPELSLILITRPTALEKHLLEHPRHVEILGFESEARKEYFYK 349
 Db 562 -----VIQKLIQKNHLSRTCLLIAVTRTPARDIRYLE--TILEIQAP-----FYN 606
 QY 350 -----YFHNAEQAGQVFNVDNEPLFTMCFVPLVCWVCT-CLQOQLEGGGLLRQT 400
 Db 607 TVCILRKLFSHNWTRLRKFMVYFGKQSLQIKQITPLFVAALCAHWFQPPD-----P 659
 QY 401 SRTTTAYMYLISL-MQPKGAPRIOPPNOBGLCSLAADGLMWNOKILFEEODLRKHGL 459
 Db 660 SFDVAVFKSYMERLSLRKATAEILKATVSSCG--ELAKGFFSCCFEFNDLDAEAGV 717
 QY 460 DGEDVSAFLNMLFQKDINCEYYSFIHLSQFQFAAMYI--LDGEGGAGPDQDQV--- 514
 Db 718 DEDEDITMCLMSKFTAQ--RLRPYRFLSFAQFELAGMLRIELDDSR---QSHQDLGLY 773
 QY 515 -----TRLUTEYAF-----SERSEFALTSRFLGFLNNESTRHL----- 548
 Db 774 HLKQINSPMVTYSAYNNFLNYVSSLPSTRAGPKIVSHLLHLVDNKESENDDVYLKH 833
 QY 549 --EKSL-----CWKVSPIHKMDL-----LOWIOSKAQ 573
 Db 834 QPEISLQMLRLGLMQLCPQAYFSVMVSEHLVLAKTAYQSNVTAAACSPFVLOFQ--- 889
 QY 574 SDGSTLOQGSLEFFSCLYEIOBEETQOALS-HFQVIVVSNIAASKMEHM-VSSFCLKRCR 631
 Db 890 --GRITULGAL---NLQYFDPHESLULRSIHFS--IRGKTSPPRAHFSVLTCTDKSQ 942
 QY 632 SAQVHLHYGATYS-----ADGED-----PARCSAGAHLLVLRPERTVL--- 671
 Db 943 VPTIDQYASAFEPNMEWERNLAEKEDNVKSYMOMQRRASPDLSGTGKWLSPKQYKIPCL 1002
 QY 672 -----LDAYSEHLAALCT---NPNIELSLYRNALGSGVKLLCOGLRHPNCKLQNL 722
 Db 1003 EVDVNDIVVGQDMLMTVFSASQRIELHLMH---SRG---FIESIR-PALELSKAS 1054
 QY 723 LKRCRISSACEDLSAA---LIANKNLTRMDLSG 753
 Db 1055 VTKCSISKL---ELSAEQELLTLPSLESLEVSQ 1086
 RESULT 11
 B97746
 hypothetical protein RC0370 [imported] - Rickettsia conorii (strain Malish 7)
 C:Species: Rickettsia conorii
 C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001
 C:Accession: B97746
 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.;
 Science 293, 2093-2098, 2001
 A>Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
 A:Reference number: A97700; MUID:21442074; PMID:11557893
 A:Accession: B97746
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-312 <KUR>
 A:Cross-references: GB:AE006914; PIDN:AAL02908.1; PID:gl5619435; GSPDB:GN00173
 C:Genetics:
 A:Gene: RC0370
 Query Match 3.3%; Score 180; DB 2; Length 312;
 Best Local Similarity 19.0%; Pred. No. 3.3e-06;
 Matches 64; Conservative 74; Mismatches 117; Indels 82; Gaps 12;
 QY 97 DQETRYDYVRKFRMLMEDNRNARLGEVNLHRY-----TRLLLVKHSNPMQ 144
 Db 34 DTEAKTIREPDKYQVQOEEDKIDNIAKSLNEENKFAILRRNLNTKSTSEIKNVEKINRMQ 93
 QY 145 VQQLLDTCGRGHARTVGHQASPIKTIETLPEPBERPEPRPYVMQAGIGKSMIAHVK 204
 Db 94 V-----DIK-----VLLGSAGIGKTLMHYLS 117
 QY 205 LQWADGKLFQGRFDYLFYINCREMOSATE--CSMODLIFSCW-----PEPSAPLOEL--- 255

Db 118 YKGEKLNWKFYVPRIRLKKELLSWTYGTNIDDEILSCFVHYCLDSNDIKLEDIKS 177
 QY 256 IRVERLLFIIDGDELKPSHPDQGPWCLCWEKRPTELLINSLIRKKLLPELSLIIT 315
 Db 178 IQDKRIILLDGGDEV-----AFISQSNRRDRDIMDSVQYK-----NIVWSS 221
 QY 316 RPTA-LEKHLRLHPRHVEITLGFSEAPKEFYFYKFINAPQAGOVNYVDNE-PLFTM 373
 Db 222 RPNVIEEMSNRFB-QKVDNTGWDSEIGIEQVHKNFYDKELGTPLKIFLDTHGQIKEI 279
 QY 374 CFVPLVCWVCTCLOQLEGGLRQTSRTTAYTML 410
 Db 280 CAVPINTALICLVMSDQ-----AIRDFQKXIVIKIL 311

RESULT 12

T52068
 RAN GTPase-activating protein 2 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52068
 R:Pay, A.; Nick, P.; Nagy, F.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25929
 A:Accession: T52068
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-545 <PAY>
 A:Cross-references: EMBL:AF214560; PIDN:AAF25948.1

Query Match 3.3%; Score 179.5; DB 2; Length 545;
 Best Local Similarity 23.4%; Pred. No. 8.3e-06;
 Matches 71; Conservative 55; Mismatches 139; Indels 39; Gaps 6;
 QY 665 RPRTVL--LDAYSEHLAAALCTNPNIELSYRNALSGVKKLLCQGLRHPCNKLQNL 722
 Db 195 RPEALEVMNIFDALQGLSSLSLSD-----NALGKGVRAFGALLKSLUS-SLEELY 248
 QY 723 LKRCRISSACEDLSAALIANKNLTRMDLSGN----- 754
 Db 249 LMDGIGSEKAAQAVBELIPSTENLRVLFHPNNMTGDEGALATAEVVKRSPLENFRCST 308
 QY 755 GVGFPGMMLLCGLRHQP-CRLQMIQLRKQCSGACQEMASVLTGNPHLVLDLTGNAL 813
 Db 309 RVGSGGIALSEALEH--CTHMEKLDLRNMFTEAGVSLSKTSLSSFKHMTLYSYLNL 366
 QY 814 EDGLRLCQGLRHPVCLRTLWLKICRLTAACDELASTLSVNQSLRELDLSLNLGDL 873
 Db 367 EDEGATAIVNALKESASPIEVLEMGNDITVEASAIAACVAKQDLKLNLSNELKDE 426
 QY 874 GVLLLCGLRHPTCKLQTLRLGICRIGSAACEGLSVYLOANHNRLRELDLSFNDLGDWGLW 933
 Db 427 GCQVIANCIIEVNSKLOQYIDMTNYYRRAGARALAHVVKKEAFKLNIDGNIISBEGIE 486
 QY 934 LLAE 937
 Db 487 ELKE 490

RESULT 13

T52063
 ran GTPase-activating protein [imported] - alfalfa
 C:Species: Medicago sativa (alfalfa)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T52063
 R:Pay, A.; Nick, P.; Nagy, F.
 submitted to the EMBL Data Library, December 1999
 A:Reference number: Z25929
 A:Accession: T52063
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-533 <PAY>

A:Cross-references: EMBL:AF215731; PIDN:AAF19528.1
 A:Experimental source: subspecies x varia; cultivar A2

Query Match 3.3%; Score 178; DB 2; Length 533;
 Best Local Similarity 22.2%; Pred. No. 1e-05;
 Matches 112; Conservative 69; Mismatches 205; Indels 118; Gaps 16;
 QY 510 PDQDVTLLIYAFSERSFLATSLRFLF---GLNNEE-----TRSHLE 549
 Db 19 PPSOSTRLM-----LVERVRNLTTPSIFSRKYGLNKDEAEKDAKETEDAAFTVASQHF 74
 QY 550 K-----SLCWKVSPIHKMDLQWIOQSKAOSDGSSTLQOGLSEFFSCLYEIQ--EEE 597
 Db 75 KEPDGDGSSAVQIYAKESKMLVLEKVGPRGKEENGELISEKGDAAVETVEDISGGRA 134
 QY 598 FI--QOALSHFQVIVVNSIASKMEHVVSSFCCLKCRSAQ--VLHLYGATYSADGEDRARC 653
 Db 135 FIDQGEASELLKPLMGPNSTFKICFSNRSGFLDAAHVVEPMLISIKDQKVEDLSD---- 190
 QY 654 SAGATLLVQLRPRPTVL--LDAYSEHLAAALCTNPNIELSLYRNALSGVKKLLCOGL 711
 Db 191 -----FIAGRPAEAEIVMNIIFSSALERAV-----LRYLNLNNAMGKGVAFRALL 238
 QY 712 RHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGVPFGMMMLLCGLRHP 771
 Db 239 KSON-DEELYLMNDGISEEAAKAVAEILPSTKLVLFHNNMTGDEGAFALAEVWKR 297
 QY 772 QCRLOMIQLRKQCSGACQEMASVLTGNPHLVLDLTGNA----- 812
 Db 298 PA-LEDFRCSSTRVSGEGGVALAEALGACTHLKLDLRDNNMFGVAGVALSKVIVFADL 356
 QY 813 -----LEDGLRLCQGLRHPCVCLRTLWLKICRLTAACDELASTLSVNQSLREL 863
 Db 357 TEIYISYNLEDDGAEALANALKESAPSELTDMAGNDITAKATYVSAECISSKQPLTKL 416
 QY 864 DLSNLNELGDLGVLLLCGLRHPTCKLQTLRLGICRIGSAACEGLSVYLOANHNRLRELD 923
 Db 417 NLSNELKDEGA-----GL-----ISKALEGRQGLSEVDLS 447
 QY 924 FNDLGDWGLMLAEG-LQHPACRL 946
 Db 448 TNLITWSGAKLIAEAVVQKPGFKL 471

RESULT 14

A45841
 T-complex-associated-testes-expressed-1 protein - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 05-Nov-1999
 C:Accession: A45841
 R:Sarvetnick, N.; Tsai, J. Y.; Fox, H.; Pilder, S. H.; Silver, L.
 Immunogenetics 31, 283-284, 1990
 A:Title: A mouse chromosome 17 gene encodes a testes-specific transcript with unusual pr
 A:Reference number: A45841
 A:Accession: A45841
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-506 <SAR>
 A:Cross-references: GB:M28821; NID:G201909; PIDN:AAA40406.1; PID:G201910

Query Match 3.2%; Score 177.5; DB 2; Length 506;
 Best Local Similarity 21.6%; Pred. No. 1.1e-05;
 Matches 93; Conservative 55; Mismatches 155; Indels 127; Gaps 11;
 QY 523 FERSFLATSLRPLFG-----LL-----NEETSHLEKSLCWKVSPIHKMDLQWIO 569
 Db 152 FFERHLENLLKLFIPGTTDPNVILDLLPLCRNVYRIHVDFL-----PVRMPTP-LQ 204
 QY 570 SKAQSDGSTLQOGLSEFFSCLYEIQEEFIQALSHFQVIVVNSIASKMEHVVSSFCCLKR 629
 Db 205 GEEQSDSGSEGEKS-----EPEKOHYQIQTLVGGGLKHLLELDLVYGVKD 248
 QY 630 CRSAQVHLHYGATYSADGEDRARCAGATLLVQLRPERTVLLDAYSEHLAAALCTNPNI 689

Db 249 CGNFWNLFY-RDVSIAATIRACHTKI-FKLRKSKVDDKARILIRSLDHPAL 306
QY 690 IELSRYNALSGVKKLCOGLRHPNCKLQNLKRCRISSACEDLSAALIANKNTMRM 749
Db 307 EELDLSHNLIGDARGAAAKLLSH----- 330
QY 750 DLGNGVGFPGMMLLCEGLRHPQCRLOMIOLRKCQLESACQEMASVLTGTPHVELDLT 809
Db 331 -----SRLRVNLANNQLOAPGAQSIALAHANTNLVFNLR 367
QY 810 GNALEDGLRLLCOGLRHPVCRILTTLWKICRTAAACDELASTLSVNOQLRELDLSLNE 869
Db 368 LNCIEDEGGOAIAHALETNKC-LSVLHLGGNKLSEPTATLSQMLTWNTTIVSLNLSCHN 426
QY 870 LGDLGVLLLCEGLRHPCTKLTQRLIGLCRIGSAACEGLSVVLOAHNHLRELDLSFNDLGD 929
Db 427 IGQDG-----GKOLLEGIS-----DNKTILEFDRLSLDVSQ 457
QY 930 WGLWLLAEG 939
Db 458 ESEYLIGQVL 467

RESULT 15

T48193

hypothetical protein F7A7.240 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T48193

R:Bevan, M.; Terry, N.; Ardiles, W.; Buyssehaert, C.; Dasseville, R.; De Clerck, R.; De

ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24487

A:Accession: T48193

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-618 <BEV>

A:Cross-references: EMBL:AL161946

A:Experimental source: cultivar Columbia; BAC clone F7A7

C:Genetics:

A:Map position: 5

A:Introns: 213/3; 255/3; 259/3; 294/3; 350/1; 376/1; 419/2; 517/3; 601/3

A:Note: F7A7.240

Query Match

Best Local Similarity 3.2%; Score 172.5; DB 2; Length 618;

Matches 122; Conservative 61; Mismatches 160; Indels 137; Gaps 28;

QY 629 RCRAQVHLHGATYSADGEDRARCAGAHFTLLVQLRPBRTVLLDAYSEHLAAALCTNPN 688
Db 123 KCNVLVEIDLSNATEMRDAD--AAVVAEARS-LERLKLGRCKWLTDMGIGCIAVGCKLN 179
QY 689 LIELSLRYNALSGVKKL---COGLR-----HPNCKLQNLR---LRC-- 726
Db 180 TVSLX-WCVGVGDVGGLLAVRCKDIRTLDLSYLPITGKCLHDILKLOHLEILLGECFG 238
QY 727 -----RISSACEDLSAALIANKNLTMDLSGNGVGFPGPMMLLC 765
Db 239 VDDDSLKSLRHDCKSLKMYKQKLDASSQNLT-----HRLTSL-LSGAGY----- 283
QY 766 EGLRHPQCKLQMIOLRKCQLESACQEMASVLTGTPHVELDITGNALDGLR---LLC 822
Db 284 -----LORLDLSHC---SSVISLDFASSLKKVSALQSIKLDGCVTPDGLKAIGTLC 332
QY 823 QGLRH-----PVCELRITML--KICRLTAACDELASTLSVNOQLRELDLSLNELDGL 873
Db 333 NSLKEVSLKCVSVTDEEAFWIGQKRL-----LEELDTDNEIDD- 374
QY 874 GVLVLLCEGLRH-PTC-KLQTLRIGLCRIGSAACEGLSVVLOAHNHLRELDLSFNDIG--D 929
Db 375 -----EGLKSISSLSLSSKLIGCL--NITDKGLSYIGMGCSNLRLELDL-YRVSIGITD 425

QY 930 WGLWLLAEGLOHPACRLQKLWLDSC-GLTAKACENLYFTLGINQTLTDL---YLTNNALG 985
Db 426 VGISTIAQGCIIH-----LETINISYCODITDKSLVSL-SKCSLLQTFESRCPCPNITSGLA 480
QY 986 DTGVRLLCKRLSHPGCK-----LRVLWLFQMDLNKMTSHSLAALRVTKPYL-DIGC 1035
Db 481 AIAVR--CKRLAKVDLKKCPSINDAGLLALAHFSONLKQINVSDTAVTEVGLLSLANIGC 538

Search completed: July 30, 2004, 13:49:54

Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:43:01 ; Search time 18 Seconds
(without alignments)
2994.034 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 5472

Sequence: 1 MLRTAGDGLCLSTYLBEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5433	99.3	1061	1 NAL12 HUMAN	P59046 homo sapien
2	2396	43.8	1034	1 C1S1 HUMAN	Q84b8 mus musculus
3	2359	43.1	1033	1 C1S1 MOUSE	Q86w24 homo sapien
4	1799	32.9	1093	1 NAL4 HUMAN	Q96mm2 homo sapien
5	1457.5	26.6	994	1 NAL4 HUMAN	Q9c000 homo sapien
6	1375	25.1	1473	1 NAL1 HUMAN	P59047 homo sapien
7	1320	24.1	1200	1 NAL5 HUMAN	Q9nx02 homo sapien
8	1309	23.9	1062	1 NAL2 HUMAN	Q8wx94 homo sapien
9	1160.5	21.3	980	1 NAL7 HUMAN	Q9rim5 mus musculus
10	1159	21.2	1111	1 NAL6 HUMAN	P59044 homo sapien
11	1115	20.4	843	1 NAL6 MOUSE	Q91ws2 mus musculus
12	1115	19.7	1033	1 NAL1 HUMAN	P59045 homo sapien
13	1078.5	19.5	854	1 NAL6 RAT	Q63035 rattus norv
14	1067	19.5	854	1 NAL6 RAT	P13489 homo sapien
15	724	13.2	460	1 RINI HUMAN	Q8h2p9 pan troglod
16	720	13.2	460	1 RINI PANTR	P29315 rattus norv
17	717	13.1	456	1 RINI RAT	P10775 sus scrofa
18	709	13.0	456	1 RINI PIG	P91w29 homo sapien
19	529	9.7	1040	1 CARP HUMAN	Q9y239 homo sapien
20	522	9.5	953	1 CAR4 HUMAN	Q8k3z0 mus musculus
21	519	9.5	1020	1 CARP MOUSE	Q8bbh0 mus musculus
22	469.5	8.6	953	1 CAR4 MOUSE	P37076 homo sapien
23	377.5	6.9	1130	1 C2TA HUMAN	P79621 mus musculus
24	317.5	5.8	1155	1 C2TA MOUSE	Q9np44 homo sapien
25	260	4.8	1024	1 BIRC HUMAN	Q9r016 mus musculus
26	200.5	3.7	1403	1 BIRC MOUSE	Q9qwk5 mus musculus
27	197	3.6	1403	1 B1R1 HUMAN	Q13075 homo sapien
28	189	3.5	1403	1 B1R1 MOUSE	Q9jib6 mus musculus
29	182.5	3.3	1403	1 B1R1 MOUSE	Q9jib3 mus musculus
30	179.5	3.3	1402	1 B1R1 MOUSE	Q9v1w3 drosophila
31	174.5	3.2	596	1 RGPI DROME	Q13066 xenopus lae
32	172.5	3.2	580	1 RGPI XENLA	Q9qk4 mus musculus
33	171	3.1	1447	1 B1R1 MOUSE	

34	161.5	3.0	589	1	RGPI MOUSE	P46061 mus musculus
35	159	2.9	1378	1	WR52_ARATH	Q9f883 arabidopsis
36	155	2.8	587	1	RGPI HUMAN	P46060 homo sapien
37	152.5	2.8	2493	1	CYAA_USTMA	P46060 ustilago ma
38	139	2.5	578	1	LR15 RAT	Q8r5m3 rattus norv
39	136.5	2.5	767	1	MEFV_MOUSE	Q9j126 mus musculus
40	134.5	2.5	906	1	RD14_ARATH	Q9x1f0 arabidopsis
41	132.5	2.4	907	1	LGR5_HUMAN	C75473 homo sapien
42	132	2.4	907	1	LGR5_MOUSE	Q9z1p4 mus musculus
43	131.5	2.4	1256	1	FLI1 DROME	Q24020 drosophila
44	130	2.4	1138	1	RD13_ARATH	Q8w3k0 arabidopsis
45	129	2.4	463	1	POF2_SCHPO	O74783 schizosacch

ALIGNMENTS

RESULT 1						
NAL12 HUMAN						
ID	NAL12 HUMAN	STANDARD;	PRT;	1061 AA.		
AC	P59046;					
DT	28-FEB-2003 (Rel. 41, Created)					
DT	10-OCT-2003 (Rel. 42, Last sequence update)					
DT	15-MAR-2004 (Rel. 43, Last annotation update)					
DE	NACHT-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like like protein 7) (Monarch-1).					
GN	NALP12 OR PYPAF7.					
OS	Homo sapiens (Human).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RP	SEQUENCE FROM N.A. (ISOFORM 1).					
RX	MEDLINE=22162427; PubMed=12019269;					
RA	Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,					
RA	Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;					
RT	"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates					
RT	activation of NF-kappa B and caspase-1-dependent cytokine					
RT	processing.";					
RL	J. Biol. Chem. 277:29874-29880(2002).					
RL	[2]					
RP	SEQUENCE FROM N.A.					
RX	MEDLINE=22451042; PubMed=12563287;					
RA	Tschopp J., Martinon F., Burns K.;					
RT	"NALPs: a novel protein family involved in inflammation.";					
RL	Nat. Rev. Mol. Cell Biol. 4:95-104(2003).					
RN	[3]					
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).					
RC	TISSUE=Lymphoma;					
RA	Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;					
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.					
RL	[4]					
RP	SEQUENCE FROM N.A. (ISOFORM 1).					
RC	TISSUE=Leukocyte;					
RX	MEDLINE=22388257; PubMed=12477932;					
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,					
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,					
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,					
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,					
RA	Diatchenko L., Marusina K., Farmer A., Rubin G.M., Hong L.,					
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,					
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,					
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,					
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,					
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,					
RA	Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,					
RA	Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,					
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,					
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,					
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,					
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalilus D.E.,					
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;					
RT	"Generation and initial analysis of more than 15,000 full-length					

DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cold autoimmune syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
 DE PYD-containing protein 3) (PYRIN-containing APAF1-like protein 1)
 DE (Angiotensin/vasopressin receptor AII/AVP-like).
 GN CIAS1 OR NALP3 OR PYPAF1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), VARIANTS FCAS MET-198; VAL-439
 RP AND GLI-627, AND VARIANT MWS VAL-352.
 RX MEDLINE=21547523; PubMed=11687797;
 RA Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A.,
 RA Kolodner R.D.;
 RT "Mutation of a new gene encoding a putative pyrin-like protein causes
 RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome.";
 RL Nat. Genet. 29:301-305(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21909508; PubMed=11786556;
 RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
 RA Mak S., Lora J.M., Briskin M., Jurman M., Cao J., DiStefano P.S.,
 RA Bertin J.;
 RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
 RT and activates NF-kB";
 RL J. Biol. Chem. 277:11570-11575(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3), VARIANT MWS MET-198, AND
 RP VARIANTS FCAS TRP-260 AND PRO-305.
 RX MEDLINE=22241334; PubMed=12355493;
 RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
 RA Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
 RA Cotter F.E., Thome M., Hitman G.A., Techopp J., McDermott M.F.;
 RT "Association of mutations in the NALP3/CIAS1/PYPAF1 gene with a broad
 RT phenotype including recurrent fever, cold sensitivity, sensorineural
 RT deafness, and AA amyloidosis.";
 RL Arthritis Rheum. 46:2445-2452(2002).
 RN [4]
 RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
 RC TISSUE=Blood;
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu J., Chen S.-J., Chen Z.;
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 RT 300 previously undefined genes expressed in CD34+ hematopoietic
 RT stem/progenitor cells.";
 RL Genome Res. 10:1546-1560(2000).
 RN [5]
 RP VARIANT FCAS MET-198, VARIANTS MWS ASN-303; MET-349; THR-439 AND
 RP ARG-569, AND VARIANT FCAS/MWS TRP-260.
 RX MEDLINE=21987640; PubMed=11992256;
 RA Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
 RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,
 RA Granel B., Frances C., Garcier F., Edery P., Boulanguet S.,
 RA Dmergues J.-P., Delpech M., Grateau G.;
 RT "New mutations of CIAS1 that are responsible for Muckle-Wells syndrome
 RT and familial cold urticaria: a novel mutation underlies both
 RT syndromes.";
 RL Am. J. Hum. Genet. 70:1498-1506(2002).
 RN [6]
 RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
 RP THR-662, AND TISSUE SPECIFICITY.
 RX MEDLINE=22062556; PubMed=12032915;
 RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
 RA Coris E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
 RT "Chronic infantile neurological cutaneous and articular syndrome is
 RT caused by mutations in CIAS1, a gene highly expressed in
 RT polymorphonuclear cells and chondrocytes.";
 RL Am. J. Hum. Genet. 71:198-203(2002).

CC -!- FUNCTION: May function as a potential inducer of apoptosis.
 CC Interacts selectively with apoptosis-associated specklike protein
 CC containing a CARD domain (ASC). This complex may function as an
 CC upstream activator of NF-kappaB signaling.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event-Alternative splicing; Named isoforms=3;
 CC Name=2;
 CC IsoId=Q96P20-1; Sequence=Displayed;
 CC Name=1;
 CC IsoId=Q96P20-2; Sequence=VSP_005520, VSP_005521;
 CC Name=3;
 CC IsoId=Q96P20-3; Sequence=VSP_005519;
 CC -!- TISSUE SPECIFICITY: Expressed in blood leukocytes. Strongly
 CC expressed in polymorphonuclear cells, undetectable or expressed
 CC at a lower magnitude in B and T lymphoblasts, respectively. High
 CC level of expression detected in chondrocytes. Low or no expression
 CC in the other tissues tested.
 CC -!- DISEASE: Defects in CIAS1 are a cause of familial cold
 CC autoinflammatory syndrome (FCAS) [MIM:120100]; commonly known as
 CC familial cold urticaria. FCAS is rare autosomal dominant systemic
 CC inflammatory disease characterized by episodes of rash,
 CC arthralgia, fever and conjunctivitis after generalized exposure to
 CC cold.
 CC -!- DISEASE: Defects in CIAS1 are a cause of Muckle-Wells syndrome
 CC (MWS) [MIM:191900]; a rare autosomal dominant fever syndrome with
 CC episodic urticaria, arthralgia, amyloidosis and progressive
 CC sensorineural deafness.
 CC -!- DISEASE: Defects in CIAS1 are the cause of chronic infantile
 CC neurologic cutaneous and articular syndrome (CINCA) [MIM:607115];
 CC also known as 'neonatal onset multisystem inflammatory disease,'
 CC or NOMID, a rare congenital inflammatory disorder characterized by
 CC a triad of neonatal onset of cutaneous symptoms, chronic
 CC meningitis, and joint manifestations with recurrent fever and
 CC inflammation.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -!- CAUTION: Ref.4 sequence differs from that shown due to frameshifts
 CC in positions 893, 918 and 926.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF410477; AAL33908.1; -;
 CC EMBL; AF427617; AAL33911.1; -;
 CC EMBL; AY051117; AAL12497.1; -;
 CC EMBL; AY051112; AAL12497.1; JOINED.
 CC EMBL; AY051113; AAL12497.1; JOINED.
 CC EMBL; AY051114; AAL12497.1; JOINED.
 CC EMBL; AY051115; AAL12497.1; JOINED.
 CC EMBL; AY051116; AAL12497.1; JOINED.
 CC EMBL; AY056059; AAL12497.1; JOINED.
 CC EMBL; AY056060; AAL12497.1; JOINED.
 CC EMBL; AY051117; AAL12498.1; -;
 CC EMBL; AY051112; AAL12498.1; JOINED.
 CC EMBL; AY051113; AAL12498.1; JOINED.
 CC EMBL; AY051114; AAL12498.1; JOINED.
 CC EMBL; AY051115; AAL12498.1; JOINED.
 CC EMBL; AY051116; AAL12498.1; JOINED.
 CC EMBL; AF420469; AAL65136.1; -;
 CC EMBL; AF468522; AAL78632.1; ALT_INIT.
 CC EMBL; AF092033; AAM14669.1; ALT_INIT.
 CC EMBL; AF418985; AAL14640.2; ALT_INIT.
 CC EMBL; AF054176; AAC39910.1; ALT_FRAME.
 CC EMBL; HGNC:16400; CIAS1.
 CC MIM; 606416; -;
 CC MIM; 120100; -;
 CC MIM; 191900; -;.

DR MIM; 607115; -. F:apoptosis activator activity; NAS.
 DR GO; GO:0016506; P:induction of apoptosis; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; NAS.
 DR GO; GO:0006954; P:inflammatory response; IMP.
 DR GO; GO:0007165; P:signal transduction; NAS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_RI; 3.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR Apoptosis; Repeat; Leucine-rich repeat; Alternative splicing;
 KW Disease mutation; Deafness.
 FT DOMAIN 1 91 DAPIN.
 FT REPEAT 218 534 NACHT.
 FT REPEAT 738 761 LRR 1.
 FT REPEAT 795 818 LRR 2.
 FT REPEAT 852 875 LRR 3.
 FT REPEAT 881 904 LRR 4.
 FT REPEAT 909 937 LRR 5.
 FT REPEAT 939 961 LRR 6.
 FT REPEAT 966 989 LRR 7.
 FT DOMAIN 688 695 POLY-GLU.
 FT VARSPLIC 718 1034 Missing (in isoform 3).
 FT VARSPLIC 719 775 /FTid=VSP_005519.
 FT VARSPLIC 834 890 Missing (in isoform 1).
 FT VARSPLIC 834 890 /FTid=VSP_005520.
 FT VARSPLIC 834 890 /FTid=VSP_005521.
 FT VARSPLIC 834 890 V -> M (IN FCAS AND MWS).
 FT VARSPLIC 834 890 R -> W (IN FCAS AND MWS).
 FT VARSPLIC 834 890 /FTid=VAR_013227.
 FT VARSPLIC 834 890 /FTid=VAR_014104.
 FT VARSPLIC 834 890 D -> N (IN CINCA AND MWS).
 FT VARSPLIC 834 890 /FTid=VAR_014105.
 FT VARSPLIC 834 890 L -> P (IN FCAS AND MWS).
 FT VARSPLIC 834 890 F -> S (IN CINCA).
 FT VARSPLIC 834 890 /FTid=VAR_014106.
 FT VARSPLIC 834 890 T -> M (IN MWS).
 FT VARSPLIC 834 890 A -> V (IN MWS).
 FT VARSPLIC 834 890 /FTid=VAR_013228.
 FT VARSPLIC 834 890 H -> R (IN CINCA).
 FT VARSPLIC 834 890 /FTid=VAR_014367.
 FT VARSPLIC 834 890 T -> N (IN CINCA).
 FT VARSPLIC 834 890 /FTid=VAR_014368.
 Query Match 43.8%; Score 2396; DB 1; Length 1034;
 Best Local Similarity 46.1%; Pred. No. 1.4e-165;
 Matches 495; Conservative 175; Mismatches 303; Indels 100; Gaps 10;

210 GKLFQGRFDYLFYINCREMNSATSCSMQDLIFSCWPEPSAPLOBLIRPERLLFIIDGF 269
 244 GTLYQDRFDYLFYIHCREVS-LVTQSLGDLIMSCCPDPNPHIKVRKPSRILFLMDGF 302
 270 DELKPSFHDPOQWCLCWEKRPTELLNSLRKLLPELSLITRTPALEKHLRLLH 329
 303 DELQAFDEHIGFLCTDOKABRGDITLLSSLRKLLPEASLITRTPALEKHLRLLH 362
 330 PRHVLILGFESEARKEYFYKYFHNAEQAGOVNYVDNEPLFTMCPVPLVWCVVCTCLOQ 389
 363 PRHVLILGFESEARKEYFYKYFSDAQAASLSIQENEVLFMCFPIPLVWCVVCTCLOQ 422
 390 QLEGGGLLQTSRTTAVVYMLYLLSIMQPKGAPRLQPPPNQORGLCSLAADGLWVKILF 449
 423 QMESKSLAQTSTTAVVYVFFLSLLQPRGSGEHLCAHLWGLCSLAADGLWVKILF 482
 450 EFQDLRKHLGDGEDVSFAFLNMIFQKINCERYYSFIHLSPOFFFAAMYIYLDDEGGAG 509
 483 EESDLRNHGLQKADVSFAFLNMIFQKINCERYYSFIHLSPOFFFAAMYIYLDDEGGAG 542
 510 -----PDODVTLTLEYAFSERSFLATSRPLFGLLNEBTRSHLSKSLCWKVSPI 560
 543 NVPGSRKLPSRDVTVLLNENYKFKGYLIFVRFPLGLVNBQRTSYLEKLSCKLSQOI 602
 561 KMDLLQWISQKASQDSTLQOGSLEFFSCLEYIEQEETIQALSHFQVIVVNSIAKMEH 620
 603 RUELLKWEVKAKAKKIQPSQLEFYLYEMQEEDFQRAMDYFPKIEI-NLSTRMDH 661
 621 MVSSPCLKCRSAQVHLVGYATYSADGEDRARCAGAHLLVQLRPERVLLDAYSEHLA 680
 662 MVSSFCIENCHRVESLSL-GFLHNPKEEEEEKEGRHDMVQ----- 703
 681 AALCTNPILIELSLYRNALSGVGLLCCQLRHPCNKLQNLRLKRCRISSACEDLSAAL 740
 704 ---CVLP-----SSHAACSHG-----LVNHLTSSFCRGLFSVL 735
 741 IANKNLTRMDLNGVGVFPMMLLCEGLRHPCQLQMIQLRKQLESGAGCAQEMASVLGTN 800
 736 STSQSLTELDLSNLSLGDPMRVLCETLQHPGNCIRRLMLGRCGLSHECCFDLSVLSN 795
 801 PHVLDELDTGNALEDGLRLCCGLRHPVCLRTLMKICRLTAACDELASTLSVNSQL 860
 796 QKVLVELDLSNALGDFGIRLLCVGLKHLNCKKULWLVSCCLTSACCQDLASVLSTSHL 855
 861 RELDLSINELGDLGVLLCEGLRHPTCKLQTLRIGICRLGSAACEGLSVVLOAHNLREL 920
 856 TRYVGENALGDSGVAILCEKAKNPQCNLQKLGIVNSGLTSVCCSALSVLSTNQNLTHL 915
 921 DLSFNDLGDWGLWLLAEGLOHPACRLQKMLDSCGLTAKACENLYFTLGINQITLDLYLT 980
 916 YLRGNTLGDGKIKLCEGLLHPDCKLQVLELDNCNLTSHCCWDLSTLTSSQSLRKL 975
 981 NNALGDTGVLLCKRLSHPCKLRLVWLFQMDLNKMTSHSLAALRVTKPYLDI 1033
 976 NNDLGDGLGVMMFCEVLKQOSCLIQNLGLSEMNYETKSALETLOEKEPILTV 1028
 RESULT 3
 C1S1_MOUSE
 ID_C1S1_MOUSE STANDARD; PRT; 1033 AA.
 AC Q8R4B8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Cold autoinflammatory syndrome 1 protein homolog (PVRIN-containing
 DE APAFL-like protein 1) (Mast cell maturation inducible protein 1).
 GN C1AS1 OR NALP3 OR PYPAF1 OR NM1G1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 [3]
 RA BERTIN J.;
 RA "PYPAF4: a novel PYRIN-containing APAF1-like protein.";
 RT Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 RA KAWAI-HIO Y., SAITO K., NISHIKAWA T., KIMURA K., YAMASHITA H.,
 RA MATEUO K., NAKAMURA Y., SEKINE M., KIKUCHI H., MURAKAWA K.,
 RA KANEHORI K., TAKAHASHI-FUJII A., OSHIMA A., SUGIYAMA A., KAWAKAMI B.,
 RA KANEHORI K., SUGANO S., NAGAHARI K., MASUHO Y., NAGAI K., ISOGAI T.;
 RT "NEDO human cDNA sequencing project.";
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RC SEQUENCE OF 436-994 FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA STRAUSBERG R.L., FEINGOLD E.A., GROUSE L.H., DERGE J.G.,
 RA KLAUSNER R.D., COLLINS F.S., WAGNER L., SHENMEN C.M., SCHULER G.D.,
 RA ALTSCHUL S.F., ZEEBORG B., BUETOW K.H., SCHAEFER C.F., BHAT N.K.,
 RA HOPKINS R.F., JORDAN H., MOORE T., MAX S.I., WANG J., HSIEH F.,
 RA DIATCHENKO L., MARUSINA K., FARMER A.A., RUBIN G.M., HONG L.,
 RA STAPLETON M., SOARES M.B., DONALD M.P., CASAVANT T.L., SCHEETZ T.E.,
 RA BROWNSTEIN M.J., USHDIN T.B., TOSHIYUKI S., CARNINCI P., PRANGE C.,
 RA RABA S.S., LOQUELLANO N.A., PETERS G.J., ABRAMSON R.D., MULLAHDY S.J.,
 RA BOSAK S.A., MCEWAN P.J., MCKERNAN K.J., MALEK J.A., GUNARATNE P.H.,
 RA RICHARDS S., WORLEY K.C., HALE S., GARCIA A.M., GAY L.J., HULYK S.W.,
 RA VILLALON D.K., MUZYNY D.M., SODERGREN E.J., LU X., GIBBS R.A.,
 RA FANEY J., HELTON E., KETTENAN M., MADAN A., RODRIGUES S., SANCHEZ A.,
 RA WHITING M., MADAN A., YOUNG A.C., SHEVCHENKO Y., BOUFFARD G.G.,
 RA BLAKESLEY R.W., TOUCHMAN J.W., GREEN E.D., DICKSON M.C.,
 RA RODRIGUEZ A.C., GRIMWOOD J., SCHMUTZ J., MYERS R.M.,
 RA BUTTERFIELD Y.S.N., KRZYWINSKI M.I., SKALSKA U., SMAILUS D.E.,
 RA SCHNERCH A., SCHEIN J.E., JONES S.J.M., MARRA M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some
 isoforms;
 CC Name=1;
 CC IsoId=Q96MN2-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q96MN2-2; Sequence=VSP_003917;
 CC Name=3;
 CC IsoId=Q96MN2-3; Sequence=VSP_003916;
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC -!- CAUTION: A stop codon in Ref.4 was read through in position 41 to
 extend the sequence.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; AF442488; AAL35293.1; -;
 DR EMBL; AY072792; AAL68396.1; -;
 DR EMBL; AF475747; AAL87104.1; -;

DR EMBL; AF482706; AAL88672.1; ALT_INIT.
 DR EMBL; AK056688; BAB71254.1; -;
 DR EMBL; BC016443; AAL16443.1; -;
 DR Genew; HGNC:22943; NALP4.
 DR InterPro; IPR007091; LRR_RNinh.
 DR Submitt (FEB-2002) to the EMBL/GenBank/DBJ databases.
 [4]
 RA KAWAI-HIO Y., SAITO K., NISHIKAWA T., KIMURA K., YAMASHITA H.,
 RA MATEUO K., NAKAMURA Y., SEKINE M., KIKUCHI H., MURAKAWA K.,
 RA KANEHORI K., TAKAHASHI-FUJII A., OSHIMA A., SUGIYAMA A., KAWAKAMI B.,
 RA KANEHORI K., SUGANO S., NAGAHARI K., MASUHO Y., NAGAI K., ISOGAI T.;
 RT "NEDO human cDNA sequencing project.";
 RN Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 [6]
 RC SEQUENCE OF 436-994 FROM N.A. (ISOFORM 2).
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA STRAUSBERG R.L., FEINGOLD E.A., GROUSE L.H., DERGE J.G.,
 RA KLAUSNER R.D., COLLINS F.S., WAGNER L., SHENMEN C.M., SCHULER G.D.,
 RA ALTSCHUL S.F., ZEEBORG B., BUETOW K.H., SCHAEFER C.F., BHAT N.K.,
 RA HOPKINS R.F., JORDAN H., MOORE T., MAX S.I., WANG J., HSIEH F.,
 RA DIATCHENKO L., MARUSINA K., FARMER A.A., RUBIN G.M., HONG L.,
 RA STAPLETON M., SOARES M.B., DONALD M.P., CASAVANT T.L., SCHEETZ T.E.,
 RA BROWNSTEIN M.J., USHDIN T.B., TOSHIYUKI S., CARNINCI P., PRANGE C.,
 RA RABA S.S., LOQUELLANO N.A., PETERS G.J., ABRAMSON R.D., MULLAHDY S.J.,
 RA BOSAK S.A., MCEWAN P.J., MCKERNAN K.J., MALEK J.A., GUNARATNE P.H.,
 RA RICHARDS S., WORLEY K.C., HALE S., GARCIA A.M., GAY L.J., HULYK S.W.,
 RA VILLALON D.K., MUZYNY D.M., SODERGREN E.J., LU X., GIBBS R.A.,
 RA FANEY J., HELTON E., KETTENAN M., MADAN A., RODRIGUES S., SANCHEZ A.,
 RA WHITING M., MADAN A., YOUNG A.C., SHEVCHENKO Y., BOUFFARD G.G.,
 RA BLAKESLEY R.W., TOUCHMAN J.W., GREEN E.D., DICKSON M.C.,
 RA RODRIGUEZ A.C., GRIMWOOD J., SCHMUTZ J., MYERS R.M.,
 RA BUTTERFIELD Y.S.N., KRZYWINSKI M.I., SKALSKA U., SMAILUS D.E.,
 RA SCHNERCH A., SCHEIN J.E., JONES S.J.M., MARRA M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Comment=Experimental confirmation may be lacking for some
 isoforms;
 CC Name=1;
 CC IsoId=Q96MN2-1; Sequence=Displayed;
 CC Name=2;
 CC IsoId=Q96MN2-2; Sequence=VSP_003917;
 CC Name=3;
 CC IsoId=Q96MN2-3; Sequence=VSP_003916;
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeats.
 CC -!- CAUTION: A stop codon in Ref.4 was read through in position 41 to
 extend the sequence.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; AF442488; AAL35293.1; -;
 DR EMBL; AY072792; AAL68396.1; -;
 DR EMBL; AF475747; AAL87104.1; -;

DR EMBL; AF229062; AAK00751.1; -;
 DR EMBL; AB023143; BAA76770.1; -;
 DR EMBL; AL117470; CAB55945.1; -;
 DR PIR; T17255; T17255.
 DR HSP; P13489; 1A4Y.
 DR MIM; 606636; -;
 DR GO; GO:0005622; C:intracellular; IC.
 DR GO; GO:0016506; F:apoptosis activator activity; NAS.
 DR GO; GO:0008656; F:caspace activator activity; NAS.
 DR GO; GO:0019899; F:enzyme binding; IPI.
 DR GO; GO:0006919; P:caspace activation; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; NAS.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 DR PROSITE; PS02029; CARD; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 92 DAPIN.
 FT DOMAIN 328 637 NACHT.
 FT REPEAT 704 725 LRR 1.
 FT REPEAT 807 830 LRR 2.
 FT REPEAT 864 887 LRR 3.
 FT REPEAT 921 944 LRR 4.
 FT REPEAT 950 973 LRR 5.
 FT REPEAT 1199 1215 LRR 6.
 FT REPEAT 1216 1236 LRR 7.
 FT DOMAIN 1374 1463 CARD.
 FT NP_BIND 334 341 ATP (POTENTIAL).
 FT VARSPPLIC 958 987 Missing (in isoform 3 and isoform 4).
 FT VARSPPLIC 1262 1305 /FTID=VSP_004326.
 FT VARSPPLIC 340 340 Missing (in isoform 2 and isoform 3).
 FT MUTAGEN 340 340 /FTid=VSP_004327.
 FT CONFLICT 155 155 K->L: NO EFFECT.
 FT CONFLICT 246 246 L->H (IN REF. 1).
 FT CONFLICT 782 782 T->S (IN REF. 1).
 FT CONFLICT 878 878 T->M (IN REF. 1).
 FT CONFLICT 995 995 T->I (IN REF. 1).
 FT CONFLICT 1119 1119 M->V (IN REF. 1).
 FT CONFLICT 1184 1184 M->V (IN REF. 1 AND 6).
 FT CONFLICT 1241 1241 V->L (IN REF. 1).
 FT CONFLICT 1366 1366 R->C (IN REF. 1).
 SQ SEQUENCE 1473 AA; 438F0DCE45C2562D CRC64;
 Query Match 25.1%; Score 1375; DB 1; Length 1473;
 Best Local Similarity 33.9%; Pred. No. 2.2e-91;
 Matches 366; Conservative 138; Mismatches 357; Indels 218; Gaps 23;
 QY 12 RLSTVLELEAVELEKFKVLGLTA--TELGEKIPWGSMEKAGPLEMAQLLITHFGPEEA 69
 DB 8 RLACYLEFLKKEELKEFQALLANKAHSRSSGETP-AQPEKTSMEVASLYVAQYGEORA 66
 QY 70 WRALSTFERINRKDLMERGQR-----EDLVRDPQETRYDYV----- 106
 DB 67 WDLALHTWEQMLRSLCAQAQEGAGHSFPYPSPPEHLGSPSQSTSTAVLMPWHELPA 126
 QY 107 -----RRKFLMRDVARLGEVCNLSHRYTRLLLVKHSNPMQVQ----- 146
 DB 127 GCTQGSERVLQLPDTSGRWRREISASLLYQALFSPDHPSPQSPESNAPTSTAVLGSW 186
 QY 147 ----- 146

DB 187 GSPQPSLAPEQEAQCTQWPLDSTSGIYYTEIREREREKSEKGRPPWAAVVGTPPQAHT 246
 QY 147 -----QQLLTGGRGHARTV----- 160
 DB 247 SLOPHHPWPSPVRESLCSTWPMKNEFDNQKFTQLLLQRPSPRSQDPLVKRSPDYVEE 306
 QY 161 --GQASPIKETLFEPEDEERPEPRVVMQAGICKSLAHKVMWLDWADGKLFQCGRED 218
 DB 307 NRHGL---IEIRDLFGGLDTQE--PRVILQGAAGIKSTLARQVKEAWGRGQLYGRFQ 362
 QY 219 YLFYINCREMNQSAATECSMODLIFSCWPSPASIQELIRVPERLLFIDGFDLKPFSHD 278
 DB 363 HVFYFSCRELAQSKV-VSLAELIGKDGCTATPAPRIQLSRPERLLFILDGVDFGWLQVE 421
 QY 279 PQGPMCLCWEEKRPTTELLNSLIRKLLPELSLLITRPTALEKHLLEHPRHVEILGF 338
 DB 422 PSELCLHWSQPQADALGSLGKTLTLPASFLITARTTALQNLIPSLQARWVEVLGF 481
 QY 339 SEAEKEVEYKYFHNAQAGOVNVRDNEPLFTMCVPLVVCVCTCLOOQLEGGLLR 398
 DB 482 SESSRKEYFYRYFTDERQAIRAFRLVSKNELWALCLVPPVWSWLACTCLMQMKRKEKIT 541
 QY 399 QTSRTTAVVMLYLLSLMQPKGAPRLQPPNQRLGCLSLAAGLWNOKILFEEODLRKHG 458
 DB 542 LTSKTTTTLCHYLAQALQALQPLGPQL-----RDLCSLAAEGIQWKTLPSPDLRKHG 595
 QY 459 LDGEDVSALNNMIPQKINCERYYSFIHLSFQFFFAAMYIILDEGEGAGPDQ----DV 514
 DB 596 LDGAIISTFLKWLQIEH-PIPLSYSTHLCFQFFFAAMSYVL-EDEKGRKHSNCIIDL 653
 QY 515 TRLLTEYAFSERSFALTSRELFGLNNEETSHLEKSLCMKVSPIHKMDLLQWISQAQS 574
 DB 654 EKTLEAYGI-HGLFGASTTRFLLGLLSDGEBRENIHFHCLSQ--GRNLMQWPSLO-- 708
 QY 575 DGSTLQOQSLEFFSCLYEIOEEETIQOALSHFQIVIVSNIASKMEHVMVSSFLKRCRQA 634
 DB 709 --LLQPHSLSLHCLVETRNKTLTQVMAHFEEMGMC-VETDMELLVCTFCIKFSRHVK 765
 QY 635 VLHLYGATYSADGBDRACSAHTLLVQLRPERTVLDDAYSEHLAAALCNPINLIEUSL 694
 DB 766 KLQLI-----EGRQHRSTWSPTMVFLR--WVPVTDAYWQILFSLVKVTRNLKELDL 815
 QY 695 YRNALSGRGVLLCQGLRHNPCKLQNLRLKRCRISSSACEDLSAALIANKLITMDLSGN 754
 DB 816 SGNLSHSAVASKLCTLRPRCKLLETTLKACGUTAECDKDLAFGLRANQITELDLDFN 875
 QY 755 GVGFPGMMLLCEGLRHPCRLQMIQLRKKCQLESACQEMASVGLTNPHLVLDLTGNALE 814
 DB 876 VLTDAGAKHLQRLRQPSCKLQRLQVSCGLTSDCCODLASVLSASPSLKELDLQONNLD 935
 QY 815 DLGLRLCQGLRHPCRLTTLWLKICRLTAACDELASTLSVNQSLRELDLSINEL----- 870
 DB 936 DVGVRLCEGLRHPCAKLIRLGL-----DQTLSDDEMRELQALQEQKPOLLIIFS 985
 QY 871 -GDGLVLLLCBGL-----RHPTCKIQLRLGICRLGSAACGLSVLQANHNRLRDLIS 923
 DB 986 RRPFSVMTPEGLDTGEMSNSTSLKQRLGSERAASHV-----AQANLKLLDVS 1035
 RESULT 7
 ID NAL5 HUMAN STANDARD; PRT; 1200 AA.
 AC PS9047;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT- LRR- and PYD-containing protein 5 (Water protein homolog).
 GN NALP5 OR WATER.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RC TISSUE-Ovary;
 RX MEDLINE=21922687; PubMed=11925379;
 RA Tong Z.-B., Bondy C.A., Zhou J., Nelson L.M.;
 RT "A human homologue of mouse Mater, a maternal effect gene essential
 for early embryonic development.";
 RL Hum. Reprod. 17:903-911(2002).
 CC -!- TISSUE SPECIFICITY: Oocyte specific.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
 CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
 Mater.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AY054986; AAL15549.1; -;
 CC Genew; HGNC:21269; NALPS.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 3.
 DR Pfam; PF05729; NACHT; 1.
 DR SMART; SM00368; LRR_R1; 11.
 DR PROSITE; PSS0824; DAPIN; 1.
 DR PROSITE; PSS0837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 57 148
 FT DOMAIN 280 602
 FT REPEAT 704 727
 FT REPEAT 730 753
 FT REPEAT 780 803
 FT REPEAT 809 832
 FT REPEAT 836 863
 FT REPEAT 865 892
 FT REPEAT 893 916
 FT REPEAT 950 973
 FT REPEAT 979 1002
 FT REPEAT 1007 1034
 FT REPEAT 1036 1059
 FT REPEAT 1064 1092
 FT REPEAT 1121 1142
 FT NP_BIND 286 293
 SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A71B28FA CRC64;
 Query Match 24.1%; Score 1320; DB 1; Length 1200;
 Best Local Similarity 32.7%; Pred. No. 1.6e-87;
 Matches 351; Conservative 167; Mismatches 461; Indels 96; Gaps 18;
 QY 17 LLEDAVELKKFK-LYLGTATLGECKIPWGSMEKAGPLEMAQLLIITHGPEANWLALS 75
 DB 69 LYLDELKEEFQTFKKKKSESTTCSIPQFIENANVECLALLLHEYYGASLAWATSIS 128
 QY 76 TPERINRDLWRGOREDLVR---DPQTYRDYVRKFKL----- 112
 DB 129 IFENMLRLTSEKA-RDDMKRHSPEPEATMTDQSPKKEKVPGISQAVQDSATAETKE 187
 QY 113 -----MEDRNARLGEVNLSHRYTKLLLVKHSNPMQVQQQLLDTRGRGHARTVGHQASP 166
 DB 188 QBISQAMEQEGATAA--TEEQEISQAMEQEGATAAETEEQ-----GHGGDTWDYKSH 238
 QY 167 IKETLTFEPDEER-----PE-----PPRTVMQGAAGIGKSMIAHK 202
 DB 239 VMTKFAEEEDVRRSFENTAADPEMOTLAGAFDSDRMGPRPRTVVLHGKSGIGKSALARR 298

QY 203 VMLDWADGKLFQGRDYLYFYINCREMNOATSCMODLIFSCWPEPSAPLOELIRVPERL 262
 DB 299 IVLWQAQGLVQGMESYVFFLPVREK-QRKKESSVTEFLSRWPDSPQAPVTEIMSRPERL 357
 QY 263 LFIIDGFDLPSFHDPOQPMCLCWEEKRPTTELLNSLRKLLPELSLIITRTALBK 322
 DB 358 LFIIDGFDL-GSVLNDTKLCKDWAQKPPPTLRSRLKVLPPESFLIVTRVDGTEK 416
 QY 323 LHRLEHPRHVEILGFSAEARKEYKYKFNHAEQAGOVNRYVDNEPLFTMCFVPLVCWV 382
 DB 417 LKSEVVSRYLLVRGIGSGQRHLLERIGISHQKTKQGLRAINNRELDQCVFVAVGSL 476
 QY 383 VCTCLOQLEGGLLRQTSRTTAVVMLYLLSLMOPKPGAPR---LQPPNORGLCSLAA 439
 DB 477 ICVALQLODVGESVAPNQTLTGLHAAFAHQLTPRGVVRCLNLEERVLKRCFMAV 536
 QY 440 DGLMNQKILFEQDRLRKHLGDGEDVSFANMNIPOKDCINERYYSFIHLSFQEFFAAMY 499
 DB 537 EGVWNRKSVFDGDDLMVQGLGESELRALPHNNILLPDSHCCEYYTFFHLSQLQDFCAALY 596
 QY 500 ILDEGEGGAGPDODVTRLLTEYAFSESP-----LALTSFLGLNNEETRSLE 549
 DB 597 VLE-----GLEIEPALCPLYVEK--TKRSMELKQAGFHLHSLMWKRLFLGNSDEVRRPLE 650
 QY 550 KSLCWKVSPIKMDLLQMTQSKAQSDGSTLOQGSLEFFSCLYEIOEEEFIOQALSHFOVI 609
 DB 651 VLLGCFVPLGVKQLLHWVSLGQQPNATTPGDTLDAFHCLFETQDKFVRLALNSFQEV 710
 QY 610 VVSNIASKMEHVMVSFCLKRCRQSAQVHLHYGATYSADGEDRARCAGAHTLVQLRPERT 669
 DB 711 WLP-INQNLDLIASFCLQHCPLYRKIRVDVKGIIPRDESAEACP-----VYPLWMDRK 763
 QY 670 VLLDAYSEHLAALCTNPNLIELSYRNALSGRVKLLCGLRHPNCKLQNLRLKRCRIS 729
 DB 764 TLIEQWEDFCSMGLGTHPLRDLGSSILTERAMKTLCAKLRHTCKIQTLMFRNAQIT 823
 QY 730 SSACEDLSAALANKLTMDLGNGVPPGMMMLCEGLRHPCQLQMTQLRKCOLLESGA 789
 DB 824 PGV-QHLWRIVMANRNLRLSLNLGGTHLKEEDVYMACEALKPKLLESRLDCGLTHAC 882
 QY 790 COEMASVLGTNPHLVELDTGNALDELGLRLLCQGLRHVPCRLRTLWLKILRLTAACDE 849
 DB 883 YLKISQILITPSLSLSLAGNKVTDQGVTPLSDALRVSCALQKLEDCGITATGCS 942
 QY 850 LASTLSVNCQSLRELDLSNELGDLGVLLCEGLRHPCQLQMTQLRKCOLLESGA 909
 DB 943 LASALVSNRSLTHLCSNNSLNEGVLNLCRSMRLPHCSLQRLMLNQCHLDTAGCGSLAL 1002
 QY 910 VILQAHNLRELDLSFNDLGDWGLWLAEGLOHPACRLQKLLDSCGLTAKACENLYFTLG 969
 DB 1003 ALMGNLSWLTHLSLMMNPVEDNGVKLLCEVMRPSCHLQDLVVKCHLTAACCELSCVIS 1062
 QY 970 INQTLTDLYTNALGDTGVRLLCKRLSHPGCKLRVLWFLMGDLNKNWTHSLAAL 1024
 DB 1063 RSRHLKSLDLTDNALDGGVAALCEGLKQ---KNSVLTRLGLKACGLTSDCCBAL 1114
 RESULT 8
 ID NAL2_HUMAN STANDARD; PRT; 1062 AA.
 AC Q9NKO2; Q9BVN5; Q9H6G6; Q9HAV9; Q9NWK3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 2 (PYRIN-containing APAF1-like
 protein 2) (Nucleotide-binding site protein 1).
 GN NALP2 OR PYPAF2 OR NBS1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]

```

or send an email to license@isb-sib.ch).
-----
CC ENBL; AF298547; ARG15253.1; ALT_INIT.
CC ENBL; AF310106; ARG30289.1; -
DR ENBL; AF464764; AAL69962.1; -
DR ENBL; AK000517; BAA91223.1; -
DR ENBL; AK025952; BAA15293.1; -
DR ENBL; AK000784; BAA91377.1; ALT_INIT.
DR ENBL; BC003592; AAH03592.1; -
DR ENBL; BC001039; AAH01039.1; -
DR HSSP; P10775; 2BNH.
DR Genew; HGNC:22948; NALP2.
DR InterPro; IPR007091; LRR_Rninh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF05729; NACHT; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW Alternative splicing.
FT DOMAIN 1 94 DAPIN.
FT FT DOMAIN 207 526 NACHT.
FT REPEAT 467 491 LRR 1.
FT REPEAT 622 645 LRR 2.
FT REPEAT 754 777 LRR 3.
FT REPEAT 810 832 LRR 4.
FT REPEAT 839 862 LRR 5.
FT REPEAT 867 890 LRR 6.
FT REPEAT 924 947 LRR 7.
FT REPEAT 981 1005 LRR 8.
FT REPEAT 1010 1033 LRR 9.
FT NP_BIND 213 220 ATP (POTENTIAL).
FT DOMAIN 518 523 POLY-GLU.
FT FT VARSPLOC 133 154 Missing (in isoform 2).
FT FT CONFLICT 1 1 M -> V (IN REF. 1).
FT CONFLICT 35 35 L -> P (IN REF. 1).
FT CONFLICT 304 304 I -> S (IN REF. 4; BAB15293).
FT CONFLICT 364 364 R -> K (IN REF. 1).
FT CONFLICT 980 980 MISSING (IN REF. 1).
FT CONFLICT 1052 1052 A -> E (IN REF. 5).
SQ SEQUENCE 1062 AA; 120514 MW; 4DB08F6E9C2BC6A7 CRC64;

Query Match 23.9%; Score 1309; DB 1; Length 1062;
Best Local Similarity 33.0%; Pred. No. 8.5e-87;
Matches 352; Conservative 176; Mismatches 430; Indels 108; Gaps 24;

Qy 13 LSTVLELEAVELKKFLYLGATELGE--GKIPWGSMEKAGPLEMAQLLITHFGEPAW 70
Db 11 LQALLEQLSQDELSSKPK-YIITTFSLAHELQKIPHEKVDKADGKQLVEILTHCDSYVWE 69
Qy 71 RLALSTERNRKDLWERGOREDLVRDPQETRYDYVRKFKRLMEDRNARLGEVCYNLSHRY 130
Db 70 MASLQVEKWHRMDSLSEKDE-----VREALKSP-----NKRKPLSLGITRKE 114
Qy 131 TRLLVKEHSPMQVQOQLDTRGHARTVGHQA---SPIK-----IETLFE----- 174
Db 115 RPLDQVDEMLERFKTEAQAFTETKGNVICLGKEVFKGKPKDKNRCRYILTKTFREMKWS 174
Qy 175 -PDEER-----PEP-PRTVVMQGAAGICKSMALHKVMLDWDACK 211
Db 175 WPGDSKEVQVMAERYKMLIPFSNPRVLPGFSPYTVVLYGPAGLGKTTIAQKMLDWAEDN 234
Qy 212 LFQGRDYLFINCREMNQATESCMDLLIFSCWPSPSAPLQELIRVPERLLFIIDGFE 271
Db 235 LIH-KFYIAFYLSCREUSRLG-PCSFRELFRDWPPELODDIPHILAAQAKLFLVIDGFE 292
Qy 272 LKPSFHDPOGPWCLCWEEKRPTPELLNSLRKLLPELSLIITRPTALEKHLRLEHPR 331
Db 293 LGAAPGALIEDICGDWEKKRPVPLVGLSLLNRVMLPKAALLVTRPRALRDILALEPEI 352
Qy 332 HVEILGTFSEARKEYFYKYPHNAEQAGQVFNVRDNEPLFTMTCFVPLVGVWVCTCLOOQL 391

```

353 YIRVEGLEEDRRAYFLRHFGDEQAMRAFAELMRNAALFOLGSAPAVCWIVCTTTLKQM 412
 392 EGGGLRLQTSRTTAVMYLLSLMQPKGAPRLOPPNORGLCSLAADGLWKNQILPEE 451
 413 EKXGDFVPTCLRTGLFLRFLCSRF---PQAGLRGA--LRTLSLIAAQLMAQTSVLHR 467
 452 QDLRKHGLGDEVDVSAFLNNIFQKDCINCRYSYFHLSPQEFPAAMYIILDSGE---G 506
 468 EDLERLGVQESDLRLFLDGIILQRVSKGVSFHLSPQFLTAIFYLEKEBEDRGG 527
 507 GAGPQDQVRLTETAFSERSFIATLSRFLGLNEETRSHLESLKWCXVSHIKMDLQ 566
 528 HTWDIGDVQKLISGVERLRNPDLIQAGYYSFGLANEKRAKELEATFGCRMSPIQKELR 587
 567 W-IQSK-AQSDGSTLOQGSLEFFSCLYEIOBEFTQOALSHPQIVVSNIAKMEHWSS 624
 588 CDISCKGGHSTVTDLQ---ELLGCLYESQEBELKVEVMAQFKEI--SLHLNADVVPSS 641
 625 FCLKCRSAQVLHLY-----GATYSADGDRARCSAGANTLLVQLRPRTVLLDAYS 676
 642 FCVKHCRLNQLKMSLQVICKENLPENVTASDSEDAEVSQDDQHWL-----PFW 688
 677 EHLAALCTNPMLIELSHYRNALSGRGVKKLQGLRHPNCKLQNLRLKRCRISSSACEDL 736
 689 TDLCSIFSGKDLMLGLAINDSFLSALVRLCEQIASDTCHLQVVVKNIS-PADAHRL 747
 737 SAALANKVLTMDLSGNGV--FPGMMLCGLRHPQRLQMIQLRKQLESAGCQENA 794
 748 CLALRGHKVTYTLTGNDQDDMFPA--LCEVLRHPECNRLRYLGLVSGSATTTQOWADLS 804
 795 SVLTGNPHIVELDTGNALDELGLRLCOGLRHPVCRLRLTLMKICRLTAACDELASTL 854
 805 LALEVNSQSTCNVLSNELLDEGAKLLYTLRHPKFLORLSLENCHLTAENCKDLAAVL 864
 855 SYNQSLRELDLSINELGDLVLLLCBGLRHPCTKLTQLRLGICRLGSAACBGLSVVLOAN 914
 865 VVSRELTHLCLAKNPIGNTVKFLCBGLRYPECKLTQLVLMNCIDTSDGCCDLTKLQEK 924
 915 HNLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGLTAKACENLYFTLGINOTL 974
 925 SLLCLDLGLNHLGKMGKFLCEALRKPLCNRLCMLWGCSPFPFSCEDLCSALSCNQL 984
 975 TDLYLTNNALGTVRLRLCKRLSHPGCKLRVLWL-----FGMDLNKM 1016
 985 VTLDLQNPFGSGGVKMLPETITCSGTLRTLRKIDDFDELNKL 1030

RESULT 9
 ID NAL7 HUMAN STANDARD; PRT; 980 AA.
 AC O8W34;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 7 (PYRIN-containing APAFI-like protein 3).
 GN NALP7 OR PYPAF3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22162427; PubMed=12019269;
 RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
 RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.,
 RT "PYPAF7, a novel PYRIN-containing Apati-like protein that regulates
 RT activation of NF-kappa B and caspase-1-dependent cytokine
 RT processing";
 RL J. Biol. Chem. 277:29874-29880 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22451042; PubMed=12563287;

RA Tschopp J., Martinon F., Burns K.;
 RT "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF464765; AAL69963.1; -;
 DR EMBL; AY154462; AAO18158.1; -;
 DR GenBank; HGNC:22947; NALP7.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW ATP-binding; Leucine-rich repeat; Repeat.
 FT DOMAIN 1 93 DAPIN.
 FT REPEAT 172 491 NACHT.
 FT REPEAT 614 638 LRR 1.
 FT REPEAT 674 697 LRR 2.
 FT REPEAT 760 784 LRR 3.
 FT REPEAT 788 810 LRR 4.
 FT REPEAT 817 840 LRR 5.
 FT REPEAT 845 868 LRR 6.
 FT REPEAT 874 897 LRR 7.
 FT REPEAT 902 928 LRR 8.
 FT REPEAT 933 957 LRR 9.
 FT NP BIND 178 185 ATP (POTENTIAL).
 SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;
 Query Match 21.3%; Score 1163; DB 1; Length 980;
 Best Local Similarity 33.8%; Pred. No. 2.9e-76;
 Matches 333; Conservative 147; Mismatches 403; Indels 102; Gaps 30;
 QY 13 LSTYLELEAEVLEKFKFLVLTATLGE--GKIPGSMKAGPLEMAQLLTHFGPEEAW 70
 DB 10 LQTLLEQNEDELKFSKLL-WAFPLEDLQKTPWSEVEADCKLAEILVN--TSSEW 66
 QY 71 --RLALSTPERINRKDMERGQEDLVDPQRTYRDYVRRKFLMEDRNARLGEVNLSH 128
 DB 67 IRNATVNIILEENLTCLCKMAKAE-MMEDGQ-----VQSIDNPGLGDAEEDSE 113
 QY 129 RYTRLLLVKEHS-----NPMQVQOQLL-----DTGRGHAR-TVCHQASPIKIEITLFP 175
 DB 114 -----LAKPGKEGWRNSMEKQSLVWKNTFMQGDIDNFHDDVTLNQ-----RFIDPLNP 163
 QY 176 DEERPEPRPTVYVMOGAAGIGKSMIAHKVMDADGKLFQGRFDYLYINCREMNQSAEC 235
 DB 164 RPRKLTPTVTVLHGPAGVGTTLAKKMLDWTDCNL-SPTLYAFYLSCKELSRMG-PC 221
 QY 236 SMQDLIFSCWPSPAPLQELIRVPERLLFIIDGFDLKPFSFHDQPGW----CLCWEKKR 291
 DB 222 SFAELISKDWPELQDDIPSLAQARILFVWDGLDELKV----PPGALLQIDCGDWEKKK 277
 QY 292 PTELLNSLIRKLLPELSLLITTPRTALEKHLRLEHPRHVEILLGFSAEKKEYKVF 351
 DB 278 PVPVLGSLKLLKMLPRAALLVTTRALRDLQLLAQQPIYVRVVEGFLSEEDRRAYFLRRF 337
 QY 352 HNAQAGQVFNVRDNEPLFTMCFVPLVCVVVCTCICQQQLGEGGLRLQTSRTTAYMYL 411
 DB 338 GDEDQAMRAFAELMRNAALFOLGSAPAVCWIVCTTTLQWKEGDDVPVCLRTGLFLRF 397
 QY 412 LLSLMQPKGAPRLOPPNORGLCSLAADGLWKNQILFEEQDLRKHGLGDEVDVSAFLNMN 471

	: : : : : : :	
398	LCSRP---POGAQURGA--LRTLSLLAAQGLWAQMSVFHREDLERLGVOESDLRFLDGD	452
472	IFQKDINCERYYSFIHLSFOEFFFFAAMYIL--DEGE---GAGAPDQDTRILLITEYAFSER	526
453	ILQRDVSKGCYSFIHLSFQQFLTALFWALEKEGEDRDGHAWDIGDVQKLLSGERLKN	512
527	SFLALTGRFPGLNEETRSHLEKSLCWKVSPHIKMIDLLQWIKQAQSDGSTLOQGSLEFF	586
513	PDLTQVGHFTFGLANAKRAKELEATFGCRMSPDIKQELLQ-CKAHLHANKPLSVTDLKEV	571
587	FSCLYEQEEBFIOALSHFOVVVNSTASKMEHWVSFCLKRCBSAQVLHYGA-----	641
572	LGCYLSEQEELAKVVPAPFKFEIS-HLTNTSEVMHCFSFKHCODLKLSLOVAKGVFL	630
642	----TYSADGEDRARCSAGAHTLVQLRPERTVLDDAYSEHLAAALC---TNPNLIBLS	693
631	ENYMDFFELDTEFE-RCT-----YLTIPNWARQDLRSRLRWTFDCSLFSSNSNLKFL	681
694	LYRNALSGRGVKLLCOGLRHPCNKONLRXKRCLSSACEDLSAALTANKNLTRMDSLGG	753
682	VKOSFLSDSSVRILIICDHVTRSTRCHLQKVEIKNV-TPDTAYRDFCUAFIGKKTLTHLTLAG	740
754	NGVFFPGMWL-LCEGLRHPQORLQMIQLRKCQLESAC-----OEMASVLGTNPHELVEL	806
741	HIEWERTMWMLCDLLRNHNKCNQLVRL-----GGHCATPEQWAEFFVVLKANQSLKHL	794
807	DLTGNALIEDLGRLLCOGLRHVPCHRLTIWLKI CBLTAACDELASTLSVNOSIRELDLS	866
795	RLSANVLLEDGAMLLYKTMTRPKHFQWLMSLENCELTEASCCKDLAAVLVVSKKLTHLCIA	854
867	LNELGDLGVLLLCGLRHPTCKLQTLRLGICRLGSAACEGLSWIQAHNHIRELDLSFND	926
855	KNPIGDGVKFCEGLSPVDCKLQTLIVLQOCSI TKLGCYFLSEAQLQECASLTNLDLSINQ	914
927	LGDWGLWLLAEGLOHPACRLQKLWL	951
915	IAR-GLWTILCOALENPCNLCXHLR	938

RESULT 10

[illegible]

```
FT VARIANT      256 256 L -> P (IN STRAIN 129/SV).
FT VARIANT      977 977 H -> Q (IN STRAIN 129/SV).
FT VARIANT     1003 1004 SS -> NN (IN STRAIN 129/SV).
SQ SEQUENCE    1111 AA; 125502 MW; 4016A5D67A1C01P4 CRC64;

Query Match
Best Local Similarity 31.4%; Pred. No. 5.2e-76;
Matches 293; Conservative 165; Mismatches 407; Indels 69; Gaps 13;

QY 162 HQASP-1K-IETLFPDEPBPPTVVMQGAAGIKGSMALHKVMDWADGKLFQGRFFY 219
Db 168 HYDSEPMKLLSDAFKP-YKTEPQPHITILHGRPGVKSAIIVLGMWAGKLFQ-KMSF 225
QY 220 LPVINCENMQATSCSMODLIFSCWPSPAPLQELIRVPERLLFTIDGDFELKSFHDP 279
Db 226 VIFFSVREIKWT-EKSSLAQLAKCEPDSWDIVTKIMSQPERLLFVIDGLDDMDSVLQHD 284
QY 280 QGPWCLCWEKRPTEILLNSLRKLLPELSLIITPTALEKHLRLHPRHEVILGFS 339
Db 285 DMTLSRDNDKEQPIVILYSLARKALLPQSFLIITRTNGLEKLSMVVSPVILVEGLS 344
QY 340 EABRKEYFYKFNHABQOVNYVDNEPLFTMCPEVPLVCVVCTCLOQOEGGGLLRQ 399
Db 345 ASRRSQVLNENSDRIQVPHSHJENHQLDQOQAPSVCSLVCEALQKKGKRCYTL 404
QY 400 TSTRITTAIV--MLYLSLMQPKPGAPRLQPPNQRCGLCLAAADGLWNOKILFEODLRK 456
Db 405 PQCTUTGLYATVIFHQLTKRPSQALSQEQITVLGLCMAAGVWTRSVFYDDDLKN 464
QY 457 HGLDGEDVSAFLNMI-FOKINCERYYSFIHLSFQEFFAAMYYIILDEGGAGDPQ--- 512
Db 465 YLSKESEITLALPHMILLQVGHNSQCYVFSHLSLQDPFAALYYVL--EGLEWNNQHFC 521
QY 513 --DVTRLLTEYAFSRSFALTSRFLPGLNENETSHLEKSLCKWKSPHMKDILLQWIOS 570
Db 522 FIENORSIMEVTRTDTRLLGKMRFLFGLMKNKDLTKLEVLFEYPIPTVEQKLQHWLSL 581
QY 571 KAQSDGSTLQGSLEFSCLEYIOEEFIQALSHFQVIVSNIAKSMHMYSSFCCLKKC 630
Db 582 IAAQVNGTSPMTLDYFLCFESQDEEFVGGALKRQFQEWLL-INQKMDLVSSYCLKHC 640
QY 631 RSAQVHLH-----YGATYSADGE----- 648
Db 641 QNLKAIKRVDIRLLSVNDTLELCPVTVQETCKPULLMWMWGNFCVSLGSLRNLKELDLG 700
QY 649 DRARCSAGHTLLVQLRPE-----RTVLDAYSEHLAAALCTNPHILIELSLYRNAL 699
Db 701 DSILSQRAMKILCLELRNQSRIQKLTFKSAEVVSGLKHLLFNSQNLKYLNLGNTFM 760
QY 700 GSRGVKLLCOGHRHNCKLQNLRLKRCRISACEDLSAALIANKNLITRMDLSGNGVGFP 759
Db 761 KDDMKLACEALKHPKCVETRLDSCELTIGYEMISTILLSTYTLKCLSLAKNRVGVK 820
QY 760 GNMJLCEGRHPQCRQMTQIARKQLESACQEMASVLTGNPHLVELDTGNALEDLGR 819
Db 821 SMISLGNALSSMCILQKILNDGCTLPASCHLLVSALFNSQNLTHLCLSNNSLGTGVQ 880
QY 820 LKCOGRHPVCRILTMLKICRLTAACDELASTLSVNSQSLRELDLSLNLGDLGVLLC 879
Db 881 QICQFLRNPECALQRLILNHCNIVDDAYGFLAWRLANNKLTHTLSLTMPPVGDGAKMLJC 940
QY 880 EGRHPTCKLQTLRLGICRIGSAACEGLSVIQAENHRLRELDLSFNDLGDGWLGLWLAEG 939
Db 941 EALKTEPTCYLELELVDCQLTQNCCEDLACMITTTTKHLKSLDLGNALGDGKVITLCEGL 1000
QY 940 QHPACKLQKLWLDSCGLTKAKCBENLYFTLGINQTLTDLTYLTNALGDTGTVRLCKKLSHP 999
Db 1001 KQSSSLRLGLGACKLTNCCCEALSLAISCNPHLSNLVKNDFTSQMLKLCSAFQCP 1060
QY 1000 GCKLRVLWLFGLMDLNKMTSRSLAALRVTKPYLDI 1033
Db 1061 VSNLGIIGLWKQYYARVRQLEVEFVPHVVI 1094
```

```
RESULT 11
NAL6 HUMAN
ID NAL6 HUMAN STANDARD; PRT; 892 AA.
AC P59044;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 6 (PYRIN-containing APAF1-like
DE protein 5).
GN NALP6 OR PYPAF5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., Distefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RL J. Biol. Chem. 277:29874-29880 (2002).
RN [2]
SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Tschopp J., Martinon F., Burns K.;
RT "NALP6: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:195-104 (2003).
RN [3]
FUNCTION.
RX MEDLINE=22275822; PubMed=12387869;
RA Grenier J.M., Wang L., Manji G.A., Huang W.-J., Al-Garawi A.,
RA Kelly R., Carlson A., Merriam S., Lora J.M., Briskin M.,
RA Distefano P.S., Bertin J.;
RT "Functional screening of five PYPAF family members identifies PYPAF5
RT as a novel regulator of NF-kappaB and caspase-1.";
RL FEBS Lett. 530:73-78 (2002).
CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
CC activation of NF-kappa-B.
CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Highly expressed in granulocytes. Detected at
CC much lower levels in T-cells.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 5 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AF479748; RAL87105.1; -
EMBL; AY154461; AAO10157.1; -
Genew; HGNC:22944; NALP6.
InterPro; IPR000767; Disease resist.
InterPro; IPR007091; LRR RNinh.
InterPro; IPR003590; LRR RNinh_sub.
InterPro; IPR007111; NACHT_NTPase.
InterPro; IPR004020; PAAD_DAPIN_dom.
Pfam; PF05729; NACHT_1.
Pfam; PF02758; PAAD_DAPIN; 1.
PRINTS; PR00364; DISEASEREST.
SMART; SM00368; LRR_R1; 3.
SMART; PS50824; DAPIN; 1.
PROSITE; PS50837; NACHT; 1.
ATP-binding; Leucine-rich repeat; Repeat.
DOMAIN 1 103
FT
```


RC TISSUE=Glial tumor;
RA Kawakami B., Sugiyama A., Takemoto M., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P59045-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P59045-2; Sequence=VSP_007068;
CC Note=No experimental confirmation available;
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 6 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY095145; AAM14632.1; -
CC EMBL; AY154466; AAO18162.1; -
CC EMBL; BC034730; AAH34730.1; -
CC EMBL; AK090621; BAC03490.1; ALT INIT.
CC Genew; HGNC:22945; NALP11.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR RNinh.
CC InterPro; IPR003590; LRR RNinh sub.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF00560; NACHT; 1.
CC Pfam; PF05729; NACHT; 1.
CC SMART; SM00368; PAAD_DAPIN; 1.
CC SMART; SM00368; LRR_R1; 9.
CC PROSITE; PS00824; DAPIN; 1.
CC PROSITE; PS00837; NACHT; 1.
CC ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
CC KW ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
CC FT DOMAIN 1 91
CC FT REPEAT 147 470
CC FT REPEAT 588 611
CC FT REPEAT 632 655
CC FT REPEAT 745 768
CC FT REPEAT 802 827
CC FT REPEAT 859 882
CC FT REPEAT 919 944
CC FT NP_BIND 153 160
CC FT VARSPPLIC 614 668
CC
CC SOMKSLVYREICSLFYTWESLRELIHFDNDLNGISERILS
CC KALEHSSCKLRILK > R (in isoform 2).
CC /FTID=VSP_007068.
CC T -> I (in REF. 4).
CC
CC CONFLICT 788 788
CC SEQUENCE 1033 AA; 117794 MW; A28880485FBA49AC CRC64;
CC
CC Query Match 19.7%; Score 1078.5; DB 1; Length 1033;
CC Best Local Similarity 28.5%; Pred. No. 4,1e-70; Indels 85; Gaps 17;
CC Matches 294; Conservative 183; Mismatches 470;
CC
CC 16 YLEEEAEVKKFLYLTATLGEKGIKPGWSEKAGPLEMAQLLTHFGPEAWRLALS 75
CC 14 YLENLSDEKFEQFKKYL--ARKILDFKLPQFLIQMTKEELANVLPISEGVQYINMLFS 71
CC 76 TPERINRDLWR--GQREDLVRDQETVYRVYRKRFRIMEDRNARLGCVNLSHRYRL 133
CC 72 IFPMRKEDLCRKIIIGRR---NRNOEACKVMRRKFMQLWE-----SHTF--- 113
CC 134 LIVKEHSPMVOQQQLLDGTRGHARTVGHQASPI--KIETLFEDEPERPEPTVMQGA 191

Db 114 -----GKHYKFRDVSDFVYLQLAYDSTSYSSANNLAVFNGE 154
Qy 192 AGIGKSMIAHKVMDWADGKLFQGRDYLFYINCREMNQSMQDLIFSCWPERPAP 251
Db 155 RASGKTTIVINLAVLRWIKGEMWQNNISYVHLTSHQ--MTNSSLABELIAKADWPDGQAP 213
Qy 252 LOELIRVPERLLFIIDGDELKPSFHDPOGQWPCWKECKRTELLLSLIRKLLPELSL 311
Db 214 IADILSDPKLLFILEDLNIREFELNYESALCSNSTQKVPFIPVLLVSLRKKWAPGWF 273
Qy 312 LITRPTALEKHLRLEHPRHVEILGFSEARKEFYKYFHNAEQAGGVNFVNRDEPLF 371
Db 274 LISSRPTGNVKTFLKEVDCCTIQLSNGKREIYFNSFFKDRQORASAAOLQIVHEDELIV 333
Qy 372 TMCVPLVWVWVCTCLOQOEGGLLRQTSRTTAVVYMLYLLSLMQPKGAPRPPPNQ 431
Db 334 GLCRVAILCWITCTVLKRQMDGRDFQLCCQCTPTDLHAFLADALTSEAGL-----TANQ 388
Qy 432 -----RGLCSIAADGLWNQKILFEEQDLRKHGLEDGVSFAFLNWNIFQKINCERYYSF 485
Db 389 YHGLLRKRLCLLAAGGLFLSTLNFSGEDLRVCVGFTEADVSVLQANILLPSNTHKDRYK 448
Qy 486 IHLGFQEPFAAMYILDEGEGAGPDQDVTRLLTEYAFSERSFLALTS--RFLGLLNEE 543
Db 449 IHLNVQEPCTAIFLM-----AVPNYLIPSGREYKREYQSDFNQVFFIFGLLAN 502
Qy 544 TRSHLEKSLCWKVSPIHMDLLQW----IQSKAOSDSTLQOQSLFFSCLYEIQEEFFI 599
Db 503 RKILETSFGYLP---MVDSEKVMYGVKMDRDEPKLTH-HMPLFYCLYENREEFV 558
Qy 600 QOALSHFOVIVVSNIAKMEHWSFCLKRCRSQAQLVHLGYATYSADGEDRARCAGAT 659
Db 559 KTIVDALMEVTV-YLQSDKDMVMSLYCLDYCCHLRTLKLSVQRIFQNKLEPIRTPASQMK 617
Qy 660 LLVQLRPERTVLLDAYSEHAAALCTNPILTELSYRNALGSRGVKLLCQGLRHPNCKLIQ 719
Db 618 SLV-----YWRICSLFYTWESLRELIHFDNDLNGISERILSKALEHSSCKLR 665
Qy 720 NLRKRCRISSACEDLSAALIANKRLTRMDLSGNGVFGPGMMLCEGLRHPQCRLOMIQ 779
Db 666 TKLSYVS-TASGFEDLLKARNRSLTYLSINCTSISLNMFSLLHDLHLEPTCQISHLS 724
Qy 780 LRKCOLESAGCOEMASVLGTNPVHLVELDTGNALDELGLRLCCGLRHPVCRLRTMLKI 839
Db 725 LMKCDLRASECEETASLLISGSGURKLTLSNPURSDGWNILCDALLHPNCTLISLVLF 784
Qy 840 CRLTAAACDELASTLSVNQSLRELDLSLNEGLDGLVLLCEGLRHPCTCKLQTLRIGICRL 899
Db 785 CCLTENCSSALGRVLLFSPTRLQDLQVNLKYNVGLVHVTFFLPFTQLEHLHSGCFF 844
Qy 900 GSAAECGLSVVLQANHNIRELDLSFNDLGDWGLMLLAEGLQHPACRLQKMLDSCGLTAK 959
Db 845 SSDICQYIAIVIAITNEKRLSLEIGSNKTEDAGMQLCGLRHPNCLMVLNIGLEECMLTSA 904
Qy 960 ACENLYFTLGINOTLTDILYTNALGDTGVRLCKRLSHPCCKLVLMFGMDLNKMTHS 1019
Db 905 CCRSLASVLTNKTLERLNLLQNLHNGDGVAKLLESISPDCVLKVGVLPLTGLNTQTQQ 964
Qy 1020 RIALLRVTKPYL 1031
Db 965 LLMTVKKRKPSL 976

RESULT 14
NAL6 RAT
ID NAL6 RAT STANDARD; PRT; 854 AA.
AC Q63035,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE PYRIN-containing APAF1-like protein 5-like (Angiotensin II/vasopressin receptor).

RINI_HUMAN
ID RINI_HUMAN STANDARD; PRT; 460 AA.
AC P13489;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental ribonuclease inhibitor (Ribonuclease/angiogenin inhibitor)
DE (RAI) (RNase inhibitor) (RI).
GN RNH OR PRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89118269; PubMed=3219362;
RA Lee F.S., Fox E.A., Zhou H.-M., Strydom D.J., Vallee B.L.;
RT "Primary structure of human placental ribonuclease inhibitor.";
RL Biochemistry 27:8545-8553(1988).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 443-461.
RX MEDLINE=89210799; PubMed=3243277;
RA Schneider R., Schneider-Scherzer E., Thurnher M., Auer B.,
RA Schweiger M.;
RT "The primary structure of human ribonuclease/angiogenin inhibitor
common repetitive module.";
RL EMBO J. 7:4151-4156(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE-Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX TISSUE-Brain, Kidney, Lymph, and Ovary;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ussdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinstead J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF COMPLEX WITH ANGIOGENIN.
RX MEDLINE=97459904; PubMed=9311977;
RA Papageorgiou A.C., Shapiro R., Acharya K.R.;
RT Molecular recognition of human angiogenin by placental ribonuclease
inhibitor -- an X-ray crystallographic study at 2.0-A resolution.";
RL EMBO J. 16:5162-5177(1997).
CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
function in the modulation of cellular activities.
CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M22414; AAA59130.1; -;
DR EMBL; X13973; CAA32151.1; -;
DR EMBL; M36717; AAA60249.1; -;
DR EMBL; AL161967; CAB82310.1; -;
DR EMBL; BC003075; AAH03075.1; -;
DR EMBL; BC011500; AAH1500.1; -;
DR EMBL; BC014629; AAH14629.1; -;
DR EMBL; BC047730; AAH47730.1; -;
DR PIR; A31858; A31858.
DR PDB; 1A4Y; 14-OCT-98.
DR Genew; HGNC:10074; RNH.
DR MIM; 173320; -;
DR GO; GO:0008428; P:ribonuclease inhibitor activity; TAS.
DR GO; GO:0006401; P:RNA catabolism; TAS.
DR GO; GO:0006401; P:RNA catabolism; TAS.
DR InterPro; IPR001611; LRR_RNinh.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR Pfam; PF00560; LRR; 3.
DR SMART; SM00368; LRR_R1; 1.
KW Repeat; Leucine-rich repeat; 3D-structure; Polymorphism.
FT INIT_MET 0 0
FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
FT REPEAT 19 47 LRR A1.
FT REPEAT 48 75 LRR B1.
FT REPEAT 76 104 LRR A2.
FT REPEAT 105 132 LRR B2.
FT REPEAT 133 161 LRR A3.
FT REPEAT 162 189 LRR B3.
FT REPEAT 190 218 LRR A4.
FT REPEAT 219 246 LRR B4.
FT REPEAT 247 275 LRR A5.
FT REPEAT 276 303 LRR B5.
FT REPEAT 304 332 LRR A6.
FT REPEAT 333 360 LRR B6.
FT REPEAT 361 389 LRR A7.
FT REPEAT 390 417 LRR B7.
FT REPEAT 418 446 LRR A8.
FT VARIANT 169 169 P -> L (in dbSNP:17585).
FT CONFLICT 422 423 /FTID=VAR_014726.
FT STRAND 2 10 RQ -> SE (IN REF. 2).
FT HELIX 16 22
FT TURN 23 28
FT STRAND 31 35
FT TURN 41 43
FT HELIX 44 52
FT TURN 53 53
FT STRAND 59 61
FT TURN 63 64
FT HELIX 68 76
FT TURN 77 79
FT TURN 82 83
FT STRAND 88 90
FT TURN 92 93
FT HELIX 98 100
FT TURN 101 110
FT TURN 112 113
FT STRAND 116 118
FT HELIX 125 136
FT TURN 137 137
FT TURN 139 140
FT STRAND 145 147
FT TURN 149 150
FT HELIX 155 157
FT TURN 169 170
FT STRAND 173 175
FT HELIX 182 195

Search completed: July 30, 2004, 13:48:14
Job time : 22 secs

```

FT STRAND 202 204
FT TURN 206 207
FT TURN 213 213
FT HELIX 214 224
FT TURN 226 227
FT STRAND 230 232
FT HELIX 239 250
FT TURN 251 251
FT TURN 253 254
FT STRAND 259 261
FT TURN 263 264
FT HELIX 269 279
FT TURN 280 281
FT TURN 283 284
FT STRAND 287 289
FT TURN 291 292
FT HELIX 296 307
FT TURN 308 308
FT TURN 310 311
FT STRAND 316 318
FT TURN 320 321
FT STRAND 325 325
FT HELIX 326 328
FT STRAND 329 338
FT STRAND 344 346
FT STRAND 351 351
FT HELIX 353 364
FT TURN 365 365
FT TURN 367 368
FT STRAND 373 375
FT TURN 377 378
FT TURN 383 385
FT HELIX 386 395
FT STRAND 401 403
FT TURN 405 406
FT HELIX 411 422
FT TURN 424 425
FT STRAND 430 432
FT TURN 434 435
FT HELIX 440 452
FT TURN 454 455
FT STRAND 457 460
SQ SEQUENCE 460 AA; 49842 MW; C3D6669E2F2BF86F CRC64;

Query Match 13.2%; Score 724; DB 1; Length 460;
Best Local Similarity 42.5%; Pred. No. 6.6e-45;
Matches 150; Conservative 64; Mismatches 139; Indels 0; Gaps 0;

QY 671 LLDAYSEHLAAALCTNPNIELSLVYNALGSRGVKLLCOGLRHPNCKLQNLRLKRCRISS 730
Db 39 LTEARCKDISSALRVNPAELNLSNELGDVGVHCVLQGLQTPSCCKIQKLSLQNCCLAG 98

QY 731 SACEIDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPQCRLOMIQIRKQLESAG 790
Db 99 AGCGVLSSTLRPTLQELHLSDNLGLDAGLQCLCEGLLDPPQCRLEKQLQLEYCSLSAASC 158

QY 791 QEMASVLTGNPHLVELDLAGNALEDGLRLCOGLRHPVCRRLTLWLKICRLTAAACDEL 850
Db 159 EPLASVLRAPKPFKELTVSNNDINEAGRVLCQGLKDSFQCLEALKLESQGVTSNCRDL 218

QY 851 ASTLSVNSQLRELDLSINELDGLGVLLCEGLRHPHTCKLQTLRLGICRLGSAACEGLSV 910
Db 219 CGIVASKASRLRELKSGNKLGDVGMALCPGLLHPSSRLRTLWIMECGITAKCGDLCRV 278

QY 911 LQANHNLRLDLSFNDLDPWGLWLLAEGLOHPACRLQKWLDSOGLTAKACENLYETIGI 970
Db 279 LRAKESLSELSLGNELGDEGARLLCETLLEPCQCLSELVKSCSFACCSHFSSVLAQ 338

QY 971 NOTLTDLTLTNALGDTGVRLLCRLSHPGCKLRVLWLFQMDLNKMTSHRLAA 1023
Db 339 NRFLLELQISNNELEDAGVRELCOGLGQPGSVLRVLWLADCDVSDSSCSLAA 391
```


Db 520 YKD - PHLTQMKCFGLNEDRVKQLERTFNCKMSLKIKSKLLQCMQVLGNSDYSPSQLG 578
 Qy 583 SLEFFSCLYEIOBEETFOALSHFQVIVVSNSTASKMEHVMVSFCLKRCRCSAQVTLHYGA- 641
 Db 579 FLEPFLCYETQDKAFISQAMRCFPKVAI-NICEKTHLVSVFCLKCHCKCLRTIRLSVTV 637
 Qy 642 -----TYSADGEDRARC-----SAGAHTLLVQLRPRTVILLDAYSEHL----- 679
 Db 638 VPEKKILKTSLPNTWDGRITHCWQDLSVLT-----NEHRELDTLHNSLNDKSAWN 691
 Qy 680 -----AAALCTNPNIIELSLYRNALGSRGVKLLCQ 709
 Db 692 ILHHELHPNCKLQKLLKFTTPDGCQDITSTSLIHNKNLMHLDLKGSIDGNGVKSCE 751
 Qy 710 GLRHPNCKLQNLRLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFGPMMLLCEGLR 769
 Db 752 ALKHPECKLQTLRLSCNLAVFCCLNIALIRSQSLFNLFTNNLLDGDGVLLCEALR 811
 Qy 770 HPQCLQMIQLRKQLESAGCOEMASVLGTPNPHLVELDLTGNALEDLGLRLLCQGLRHPV 829
 Db 812 HPKCYLERLSLSCGLTEAGCEYLISLALISNKRLTHLCIADNLVDGCVKLMDSALQHAQ 871
 Qy 830 CLRRTLWLKICRLTAAACDELASTLSVNSQSLRELDLSNELGDLGVLLCEGLRHPTCKL 889
 Db 872 CTLKSLVLRCHFTSLSSEYLSLTHLSDGNSLQDNGVVKLLCDVFRHPSCNL 931
 Qy 890 QTLRLGICRLGSAACEGLSVVLQANHNIRELDLSFNDLGDWGLWGLAEGLOHPACRLQKL 949
 Db 932 QDLMLGCVLTNACCLDIASVILNPNRLSLDLGNLDQDGDGVKILCDALRYFNCNQLRL 991
 Qy 950 WLDSCGLTAKACENLYFTLGINOTLLTDLTNNAALGDTGVRLLCCKLSHPGCKLRVLMLF 1009
 Db 992 GLEYGCLTSLCCQDLSLALCNKRLIKNLTQNTILYEGIVKLYKLSKPCKLQVLGLC 1051
 Qy 1010 GMDLNMKMTSHRLAALRVTKPYLDI 1033
 Db 1052 KEAFDEEAQKLEAVGVSNPHLII 1075
 RESULT 3
 Q86W27 PRELIMINARY; PRT; 986 AA.
 ID Q86W27
 AC Q86W27
 DT 01-JUN-2003 (TremBLrel. 24, Created)
 DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
 DE NALP9.
 GN NALP9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2451042; PubMed=12563287;
 RA Tschopp J., Martinon F., Burns K.;
 RT "NALP9: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
 DR EMBL; A154464; AAC18160.1; -
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005840; C:ribosome; IEA.
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
 DR GO; GO:0005529; F:sugar binding; IEA.
 DR GO; GO:0006412; P:protein biosynthesis; IEA.
 DR InterPro; IPR001304; Lectin C.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh sub.
 DR InterPro; IPR007111; NACHT NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN dom.
 DR InterPro; IPR001865; Ribosomal_S2.
 DR Pfam; PF02758; PAAD DAPIN; 1.
 DR SMART; SM00368; LRR RI; 10.

DR	PROSITE; PS00615; C TYPE LECTIN_1; 1.	
DR	PROSITE; PS50824; DAPIN; 1.	
DR	PROSITE; PS50837; NACHT; 1.	
DR	PROSITE; PS00962; RIBOSOMAL_S2_1; 1.	
DR	SEQUENCE 986 AA; 112592 NW; B424846P70413B2 CRC64;	
QY	Query Match 27.5%; Score 1505.5; DB 4; Length 986;	
QY	Best Local Similarity 34.2%; Pred. No. 1.2e-121;	
QY	Matches 344; Conservative 193; Mismatches 383; Indels 85; Gaps 18;	
QY	16 YLEBEAELVKPKFLYIGTATELGEGK-IPWGSMEKAGPLEMAQLLTIHFGPBEARLAL 74	
QY	14 YLEKLRKEEWFKEKELLKQPLEKFKLPWAEELKKAASKEDVAKLDDKHYPGQAEVTL 73	
QY	75 STEFIRNRKDLWEGQSEDIVRQDETIRYDVRKRFKLMEDRNARLGEVCNLSHRYTRLL 134	
QY	74 NLFQINRKDLWTAKQEE--MRNKLNPYRKHKMETQLIWEKET---CLHVPEHFYKET 127	
QY	135 LVKESHPMQVQOOLLDTGRGHARTVGHQASPIKETLPEPDEERPEPTTVVMQGAAGI 194	
QY	128 MKNEY-----KELNDAYTAAR-----HTVLEGPDGI 156	
QY	195 GKSLAHKWLMDWADGKLFQGRFDYLFYINCREMNOQATECSMODLIFSCWPPSPAPLOE 254	
QY	157 GKTLLIRKWLMDWAEGLNWKDRFFVFLNVCEANGIA-ETSLLELLSRDWPSESSEKIED 215	
QY	255 LIRVPERLLFTIDGFD-----ELKPSFHPDQGGWMLCWEKRPTELLLSLTKLLPE 308	
QY	216 IFSQPERILFIMDGFEQKFNQLKADLSDD-----WRQPMPIILSSLLQKMWLPE 269	
QY	309 LSLITRPTALEKHLRHLLEHPRHVEILGSEAEKKEFYKYFHNASQAGOVFNVVRDNE 368	
QY	269 SSSLIALGKLAQMHYFMLRHPKLIKLLGSESEKKSIFYSVFGEKSKALKVFNVRDNG 328	
QY	369 PLFTWCFFVPLVCVVCTCLQOOLGEGGLRRTSTTAVMYMLYLLSLMQPKGAPRIOPP 428	
QY	329 PLFILCHNPFTCNLVCTCVKQRLBERGDELINSQNTIYVASFLLTVF--KAGSQSPPK 386	
QY	429 PNO---RGLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSAFLNNNIFQKIDINERYYSF 485	
QY	387 VNRARLKSICALAEGSIWYTFVFSHGDLRRNGLSESGVMVMGRLLOQRGDC---FAF 443	
QY	486 IHLSPQEFPAAMYIILDBEGGAGPD-QDVTRLLTEYAFSERSFLATLSRFLGLNNEET 544	
QY	444 MHLCIQEFCAAFYLLRKPDDPNPAIGSITOLVRASVVQPTLLTQVGIWMFISTEEI 503	
QY	545 RSHLEKSLCWKYSPIHKMDLLQWIOSKASQSGSTLQOGLSEFFSFCLYIEQEEFTQQALS 604	
QY	504 VSMLETSPGFLUSKDLQBEITQCLSELSQCEADREAIQAFQELFGLFETQEKFEFTVKWN 563	
QY	605 HFO--VIVVSNIASKMEHMV--SSFCIKRCSRSAQVILHYGATYSAGEDEARCSAGHTLL 661	
QY	564 FFEVFIYIGNI---EHLVIAFSLKHCQHLTLTRMCVENIFPDD--SGCISDYNEL 616	
QY	662 VOLRPRTVLLDAYSEHLAAALCTPNLIELSLRYNALGSRGVKLLCQGLRHNPCKLQWL 721	
QY	617 V-----YWRLECSMFITKNFQILDMENTSLLDPSLAILCKALAPVCKLRKL 664	
QY	722 RLKRCRISSACEDLSAALLIANKNITRMDLSNGVGVFGPMCLLCEGLRHPCQRLQMIQR 781	
QY	665 IFTSVYFGHDS--ELPKAVLHNPHLKLSLYGTSLSQSDIRHLCEYLKHPMCKIEBLIG 722	
QY	782 KCOLSEGAQOEMASVLGTNPHLVELDLTGNALEDGLRLLLCOGLRHVPVRLTFLWKIKR 841	
QY	723 KCDISEVCEDIASVLACNSKLSLNVENPLRDEGMTLLCEALKHSHCALERLMLMYCC 782	
QY	842 LTAAACDEIASTLSVNOQSIRELDLSNELGDGLVLLCEGLRHPTCKLQTLRLGICRLGS 901	
QY	783 LTSVSCDSISEVLLCSKSLSLDLGNSNALEDNGVASLCAALHKPGCSIRELWMGCFLTS 842	
QY	902 AACEGLSVVLQANHNLRELDLSFNDLDGMCLMWLLAEGLQHPACRLQKLMDSOGLTKAK 961	
QY	843 DSKCDIAAVLICNGKUKTLKGHNEIGDTGVQQLCAALQHPCKLECLGLQTCPIRACC 902	

```
QY 962 ENIYFTLGINOTLDTLVTNNALGDTGVRLCKRLSHPCCKLRVL 1006
Db 903 DDIAAALIACTLASLNDLWALDADAVVLCALSHLDFDCAQL 947

RESULT 4
Q86W87 PRELIMINARY; PRT; 994 AA.
AC Q86W87;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Strausberg R.;
RL Submitted (Apr-2003) to the EMBL/GenBank/DBJ databases.
RE EMBL: BC050326; AAH50326.1; -
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF02758; PAAD_DAPIN_1.
DR SMART; SM00368; LRR_RI; 8.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
KW Hypothetical protein.
SQ SEQUENCE 994 AA; 113434 MW; 70B76FF836AC1E68 CRC64;

Query Match 26.5%; Score 1450.5; DB 4; Length 994;
Best Local Similarity 33.8%; Pred. No. 7.7e-117;
Matches 352; Conservative 197; Mismatches 403; Indels 89; Gaps 19;

QY 13 LSTYLEBAVELKFKVLGTAT-ELGEGKIPWGSMEKAGPLEMAQLLTHPGPREAWR 71
Db 11 LMYLEELKEEFKEKHEKHLKQNTLQELKQIPWTEVKASRELANLLIKHYEQQAN 70

QY 72 LALSTERINRDKLWREGQREDIVRDPQETRYDYVRKPF-RLMEDRNARLGEVNLSHRY 130
Db 71 ITRIFQKMDRDLCKMKNRER--TGTYKTYQAHAKQTSRLWSSKSV----- 116

QY 131 TRLLLVKHSNPMVOQQLDTCGRGHARTVGHQASPIKTIETLFEPPDEERPEPTVMQG 190
Db 117 TEHLVFEF-----EVKQESCD-----HLDRLPAP-KEAGKQPRVTVIQQ 155

QY 191 AAGIGKSMIAHKVMDLWADGKLFQGRFDYLYFINCREMNSQATECSMDLITSCWPEPSA 250
Db 156 PQSIGKTLMLKMAWSNKFIDRFELTYFFCCHELRB-LPPTSIALDLISREWDPA 214

QY 251 PLOELIRVPERLLFIIDGDELKPSHPDQGGPWCWCEKRPTELLNSLIRKLLPELS 310
Db 215 PITEIVSQPERLLFVIDSFEEQGLNEPDSLCGDLMEKRPVQVLLSLRLKMKLPEAS 274

QY 311 LLITRPTALEKHLLEHPRHVEILGFSEAEKRYFYKHNAEAGQVFNVRDNEPI 370
Db 275 LLIAIKPVCPELRQVTVISEIYQPGFNESDLVYFCFFKDPKRAEAFNLVRESEQL 334

QY 371 FTMCFVPLVGVVCTCLOQLEGGLRQTSRTTAVVYMLYLSLMQPKPGA--PRLQPP 428
Db 335 FSICQIPLLCWLCTSLQEMQKGDALTCQSTTSVYSSFFVFNFTP-DGAEPTQQQTQ 393

QY 429 PNQRGICSLAAGLNQKTLFEQDLKRGHLDGEDVSAPLANWNIPOKINDICERYSFHIL 488
Db 394 HOLKALCSLAAGMWTDTFECEDLLRRNGVVDADIPALLGKTLKLLKYGERESSYVFLHV 453

QY 489 SFQEFFAAMYI-----LDEGEGAGPDQDQVTRLLTEYAFSERSFLALTSRFLGLNNEET 544
Db 489 SFQEFFAAMYI-----LDEGEGAGPDQDQVTRLLTEYAFSERSFLALTSRFLGLNNEET 544

454 CIOEFCAALFYLLKSHLDHPHAPVRCVQEL--LVANFEKARRAHWIFLGCELTGLLNKKE 511
545 RSHLEKSLCWKVSPIHKMDLLQWTSQAQSGDSTLQOQSLEFFCSCLVEIOBEFFIOQALS 604
512 QEKLDAPFGFQLSQEIQQIHCQLKSLGERGNPQGVDSLAIFYCLFEMQDPAPVQKAVN 571
605 HFQVI---VVSNIASKMEHVVSPFLKRCRQAQVHLHYGATYSADGEDRARCAGATILL 661
572 LLQEAAPHIIDNV-----DLVVSAYCLRYCSSRLKLCF--SVQNVFKKEDHSSTSDSYLI 625
662 VQLRPERTVLDDAYSEHLAAALCTNPNIIEISLYRNALGSRGVKLLCQGLRHPNCKLQNL 721
626 CW-----HHICSVLTTSGLHRELQVQDSTLSESTFTVWCNQLRHPSCRLOKL 672
722 RLKRCRISSAC-----EDLSAALIANKNITRMDLSGNGVGFPMMLLCCGLRHPQ 772
673 GINNVSFGQSVILLFEVLFYQPLDKYLSFTLTKLSRDDIRS-----LCDALNYP 722
773 CRLQMIQRLKQLESAGACQEMASVIGTNPVHVELDITGNALDELGLRLCQGLRHPVCR 832
723 GNVKELALVNVCHSLSPDCEVLAGLLTNKKLTLYLVNSCQL-DTGVFLLCEAMCSPDTVL 781
833 RTLWLKICRLTAACADELASTLSVNSQSLRELDLSNBLGDLGVLLLCGLRHPPTCKLQTL 892
782 VYLMALAFCHLSEQCCEYISEMLLRNKSRYVLDLSANVLKDEGLKTLCEALKHPDCCLDL 841
893 RLGICRLGSAACGLSVVLOANHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLWD 952
842 CLVKCFITAAAGCEDLASALISNQNLKLTQICNEIGDVGVQLLCRALTHTDCLRILEGL 901
953 SCGLTAKACENLYFTLGINOTLDTLVTNNALGDTGVRLCKRLSHPCCKLRVLWFGMD 1012
902 ECGLTSTCKDXLASVLTCKTLQLLNLTLNTLHTGVVVLCEALRHPHPECALQVLGRKTD 961

1013 LNRQTHSRLLAALRVTKPYLDI 1033
962 FDETOALLTAEBERNPLTI 982

RESULT 5
Q86W25 PRELIMINARY; PRT; 1043 AA.
AC Q86W25;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NALP13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Teschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
DR EMBL; AY154468; AA018164.1; -
DR GO; GO:0005351; P:sugar porter activity; IEA.
DR GO; GO:0009403; P:phosphoenolpyruvate-dependent sugar phospho...; IEA.
DR InterPro; IPR002114; HPr_SerP_S.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh_sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_RI; 9.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
DR PROSITE; PS00589; PTS_HPR_SER; 1.
```

SQ SEQUENCE 1043 AA; 118910 MW; C321FBB6C5206DFF CRC64;
 Query Match 25.6%; Score 1401; DB 4; Length 1043;
 Best Local Similarity 32.8%; Pred. No. 1.7e-112;
 Matches 357; Conservative 171; Mismatches 405; Indels 154; Gaps 22;
 13 LSTYLELEAVELEKFKLYL-----GTATLGEKIPWGSMEKAGPLEMAQLLTHF 64
 17 LLPYLMALDQLEBEFKLEPPQQLMDFWSPQGHFRIPWANRAADPLNLSFLDEHF 76
 65 GPBEAWRLALSTPERINKDLWRCQ-----REDL----- 94
 77 PKGQAKVVLGIFQTMNLTSLCEKVAEMKENVQLEQDPTQEDLEMLEAAGNMQTQ 136
 95 VRDQETVRDVV-----RRKFLMEDRNARLGEV-VN 125
 137 CODPNEQELDEBETGNVQAGQDPNQEPEMLEEADHRRKYR--ENMKAELLETDWN 194
 126 LS----HRYTRLLAVKEHSPMVQOQLDTRGHARTVGHQASPIKIETLPEPDEEPE 181
 195 ISWPKDHVYIRNTSKDEH---EELQRLDLPNRTAQA----- 228
 182 PPRIVMGAAGIGKSMIAHKVMDWADGKLFQGRFDYLVYINCREMNSATECSMODLI 241
 229 --QTVLVGRAGVGKTTLAMRAMLHWANGVLVQQRFYSVFLSCHKI-RYMKETTFABLI 285
 242 FSCWPEPSAPLOQLRVPERLFIIDGDEL-----KPSFHDPOGPMCLCWEKRPTEL 295
 286 SLDWDFDAPTEEFMSQEKLLFIIDGFEELIISRSRSESLD--GSPCTDWYQELPVTK 343
 296 LINSIRKKLPELSLITRTALEKHLRLEHPRHVEILQPSAERKEYKYKFNHAE 355
 344 ILHSLKLELPLATLLTIKTWFVRDLKASLVNFCVQITGFTDGLRVYFMRHFDSS 403
 356 QAGQVFNVRDNEPLTFWCFVPLVNCVWCTCLQQOLEGGGLLRQTSRTTAVYMLYLSL 415
 404 EYEKILQLRKNETLPHSCAPMVCWVCSLQPKRVYDIQSTQITTSIYAFNSL 463
 416 MQPK-----PGAPRLQPPNORGLCSLAADGLNOKILFEBQDQKHLGDGEDVSA 466
 464 FSTAEDVLADDSWFG-----QWRALCSLAIEGLMSNFTFNKEDTEIEGLEVPFIDS 515
 467 FLNNMIFOKINCERYYSFIHSFQFFFAAMYVILDEGE---GGAPDQDVTELLTEYAF 523
 516 LYEFNLIQINDCGGCTTTHLSFQFFFAAMSVLEEPREFPHSTPKQEMKMLQHVLL 575
 524 SERSFALTSRLFGLLNETRSHLEKSLCWKVSPIKMDLLOWIQSKAQSGSTLQOQS 583
 576 DKEAYWTPVVLFFGLLNKNIARELEDTLHCKISPRVMEELLKWGELCKAESASIQFHI 635
 584 LBFFSCLYEIOBEFTQALSH-POVIVVSNIAASKMEHVSFCLKRCRSQAQVHLHYGAT 642
 636 LRLFHCHESQEDFTKMLGRIFEDVL--NILEDELOQASSFCLAKHCKRL----- 684
 643 YSADGEDRARCAGATLLVLQRPVTEVLLDAYS---EHLAALCTNPNIELSLYNAL 699
 685 -----NKLRLSVSHILREDLLEITETSKPDSRMHAWNSICSTLVNENHELDLSNKL 738
 700 GSRGVKLLCOGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGF 759
 739 HASSVGLCALNPKRCVKQLTKSV-TPFWLQDLIIALQNSKLTHLNFSNKLGMT 797
 760 GNMLLCEGLRHPQCRLOMIQLRKQLESAGCAQEMASVLGNPHLVELDLTGNALEDGLR 819
 798 -VPLILKALHSAACNLKYLCKEKNLSAASQDLALFLTSIQHVTRICLGFNRLQDDGIK 856
 820 LLCGLRHPVCRRLTLMKICRLTAACDELASTLSVNSQSLRELDLSINELGDLVLLC 879
 857 LLCAALTHPKCALERLELWFCQLAAPACKHLSDALQLNRLSLTHLNSKSLRDEGVKFLC 916
 880 EGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQANHLRELDLSFNDLGDWGLMLLAEG 939
 917 EALGRPDGNLQSLNLSGCSFTREGCGGELANALSHNVKILDLGENDLQDDGVKLLCEAL 976

Qy 940 QHPACRLQKLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRLSHP 999
 Db 977 K-PHRAHTLTLGLAKCNLTATTACQHLFSLVSSSLVNLNLGNELDTDGVKMLCKALKKS 1035
 Qy 1000 GCKLRVL 1006
 Db 1036 TCRLQKL 1042
 RESULT 6
 Q86UB5 PRELIMINARY; PRT; 1375 AA.
 AC Q86UB5;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to death effector filament-forming Ced-4-like apoptosis protein.
 DE Homo sapiens (Human)
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RX MEDLINE=2388257; PubMed=12477932;
 RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ussidi T.B., Ioshizuka S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.C., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Blood;
 RA Strausberg R.;
 RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC051787; AAH51787.1; -;
 DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
 DR GO; GO:0006952; P:defense response; IEA.
 DR InterPro; IPR000767; Disease_resist.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR PRINTS; PR00364; DISEASERSIST.
 DR SMART; SM00368; LRR_RI; 5.
 DR PROSITE; PS00824; DAPIN; 1.
 DR PROSITE; PS00837; NACHT; 1.
 DR PROSITE 1375 AA; 154969 MW; 7C834D47BBD490FE CRC64;
 SQ SEQUENCE 1375 AA; 154969 MW; 7C834D47BBD490FE CRC64;
 Query Match 25.2%; Score 1377; DB 4; Length 1375;
 Best Local Similarity 33.9%; Pred. No. 3.1e-110;
 Matches 366; Conservative 138; Mismatches 357; Indels 218; Gaps 23;

Db 597 VLE---GLEIPALCPYVEK--TKSMELKQAGFHHISLMKRFELGLVSEDDVRPLE 650
Qy 550 KSLCWKVSPIKMDLLOWIKQAQSDGTLQOQSLEFFSCYELQEDEFIQQALSHFQV1 609
Db 651 VLLGCPVPLGVKQKLLHWSLLGQOPNATTPGDTLDFAHCLFQDKEFVRLALNSFQEV 710
Qy 610 VVSNIAKMBHMYSSFCXKRC-----RSAQVL----- 636
Db 711 WLP-INQNLDIASSFCLOHCPYLKIRVDVKGIFPRDESAEACPVVPLWMRDKTLIEEQ 769
Qy 637 -----HL-----YGATYSADGEDRARCAGAH-TLLVQLRPERVTLDAYSEHL 679
Db 770 WEDPCSMGLTHPHLRQLDGLSSILTERAMKTLCAKLRHPTCKIOTLMFRNAQITPGVQHL 829
Qy 680 AAALCTPNLELSIXRNALSGRGVLLCCOGLRHPNCKLQNLRLKRCRISSSACEDLSAA 739
Db 830 WRIVANRNRLSLNGLGTHLKEEDVRMACEALKHPKCLLESRLDCCGLTHACYLKISQI 889
Qy 740 LIANKNLTRMDLSNGVGFPGMMLLCEGLRHPQCRLOQMIOQRKQLESAGCAQEMASVLTG 799
Db 890 LTTSPSLKSLSLAGNKVTDQGVMLPSDALRVSCALQKLILEDGCGITATGCGQSLASALVS 949
Qy 800 NPHVELDLTGNALDELGLRLCCOGLRHPVCELRTLMLKICELTAACDELASTLSVNOS 859
Db 950 NRSILHLCLSNLSLNEGVLNLCRSMRLPHCSLQRLMLNQCCHLDTAGCGFLALALMGNWS 1009
Qy 860 LRELDLSNELGDLGVLLCEGLRHPCTCKIOTLRIGICRLGSAACEGLSVVVLQANHNLR 919
Db 1010 LTHLSLWNPVEDNGVKLLCEVWRPESCHLODELVKCHLTACCESLSCVISRSRHLKS 1069
Qy 920 LDLSFNDLGDWGLWLLAEGLQHPACRLQKMLDSCGLTAKACENIYFTLGINOTLTDLYL 979
Db 1070 LDLTNLDGCVGAALCEGLKQKNSVLTGLKAGCLTSDCCEALSLALSCNRHLTSLNL 1129
Qy 980 TNNALGDTGVRLLCKRLSHPGCKRLVL-----WLFMDLNKMTSHSLAALRVTKP 1029
Db 1130 VONNFPKGMKLCFAFACFTSNLQIIGLWKQYVPQVIRKU-----LEEVLQKLP 1179

RESULT 8
Q9BY26 PRELIMINARY; PRT: 287 AA.
AC Q9BY26;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Leucine-rich-repeat protein RNO2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
RP Shami P.J., Kanai N., Wang L.Y., Vreeke T.M., Parker C.J.;
RA "Identification and characterization of a novel gene that is
RT "Upregulated in leukemia cells by nitric oxide."
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF231021; AAK14942.1; -
DR HSP; P13489; I44Y.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR Pfam; PF00560; LRR; 1.
SQ SEQUENCE 287 AA; 31765 MW; BD3816C3255B2F9E CRC64;

Query Match 24.0%; Score 1314.5; DB 4; Length 287;
Best Local Similarity 80.1%; Pred. No. 8.4e-106;
Matches 261; Conservative 1; Mismatches 59; Gaps 2;
Indels 59; Gaps 2;
Qy 710 GLRHPNCKLQNLRLKCRISACEDLSAALIANKNLTRMDLSNGVGFPGMMLLCEGLR 769
Db 21 GLLQF--RRQLWLKRCRISSACEDLSAALIANKNLTRMDLSNGVGFPGMMLLCEGLR 78
Qy 770 HPQCRLOMIQRKQLESAGCAQEMASVLTGNPHLVLDLTGNALDELGLRLCCOGLRHPV 829

Db 79 HPQCRLOMIQRKQLESAGCAQEMASVLTGNPHLVLDLTGNALDELGLRLCCOGLRHPV 138
Qy 830 CRRLTLMWKICRUTAAACDELASTLSVNQSLRELDLSNELGDLGVLLCEGLRHPCTCKL 889
Db 139 CRRLTLMWKICRUTAAACDELASTLSVNQSLRELDLSNELGDLGVLLCEGLRHPCTCKL 198
Qy 890 QTLRLGICRLGSAACEGLSVVVLQANHNLRDLDSFNDLGDWGLWLLAEGLQHPACRLQKL 949
Db 199 QTLR----- 202
Qy 950 WLDSGGLTAKACENIYFTLGINOTLTDLYLTNNALGDTGVRLLCKRLSHPGCKRLVWLWF 1009
Db 203 -LDSCGLTAKACENIYFTLGINOTLTDLYLTNNALGDTGVRLLCKRLSHPGCKRLVWLWF 261
Qy 1010 GMDLNKMTSHSLAALRVTKPYLDIGC 1035
Db 262 GMDLNKMTSHSLAALRVTKPYLDIGC 287

RESULT 9
Q8BU40 PRELIMINARY; PRT: 982 AA.
AC Q8BU40;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to PAN2 protein.
GS E330028A19RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Ovary;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573 (2002).
DR EMBL; AK087843; E330028A19RIK.
DR MGI; MGI:2443697; E330028A19RIK.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 982 AA; 112600 MW; 6E2DBCFEA2053585 CRC64;

Query Match 23.0%; Score 1259; DB 11; Length 982;
Best Local Similarity 31.0%; Pred. No. 3.6e-100;
Matches 325; Conservative 202; Mismatches 415; Indels 106; Gaps 24;
Qy 13 LSTYLEELAEVLEKFKKLYL-GTATBELGEGKIIPWSMEKAGPLEMAQLLITHEGPEAWR 71
Db 10 LMWYLEELNKKFVKFKFLKQVQLGLKQVSWTEVKKASQDQLASLLKHYEEKPAWD 69
Qy 72 LALSTERINRKLWEGQEDLVDPQETRYDYRVRKFKRLMEDRNARLCECVNLSHRYT 131
Db 70 MTRFFQKINRKLIERAKE--IDGCPKLYRAHMTK-----MTDSS 111
Qy 132 RLLLVKSHSNPMVOQQQLDTRGHARTVGHQASPIK-----KFTEDDYDCFNLFQ-SKGTESKP 183
Db 112 RAFTI-----SIQNFKE-----KFTEDDYDCFNLFQ-SKGTESKP 147
Qy 184 RTVMQGAAGIKGSMIAHKNVMDWADOKLFGQRPDYLYFYINCREMNGSATECSMDLIFS 243

Db 148 QVFLSGAGVGKTLMLKRLMLAWIESVFLHFKFSYIFFCREVKQLKT-ASLAELISR 206
Qy 244 CWPEPSAPLOELIRVPERLFIIDGDELKPSFHDPOGPMWCLWEEKRPTPELLNSLRK 303
Db 207 EWGSPAPIEILSKPKLLFIIDSLGMECDLFWKSELCDNTEKQPVNVLSSLLR 266
Qy 304 KLPELSLLITTRPTALEKHLRLLEHPRHVEIL-GFSEAEKKEYFYKFFHNAEQGVFN 362
Db 267 KMLPESSLLISATSEPEKEMNREY-THVKIKGLKERNIKMSFHLRFDQNEAHEAFS 325
Qy 363 YVRDNEPLFMWCFVPLVWVCTCLOOQEGGLLROTSRTTAVMYLLSLMQPFG- 421
Db 326 LVRENEQLFVQVPLVWVCTCLOOQEGGLLROTSRTTAVMYLLSLMQPFG- 385
Qy 422 APRLOPPNORGLCSLAADGLWNLKILFEEDQRLKHGLDGEDVSFAFLNMFQDINCR 481
Db 386 SPKSKSDQQLGLCSLAAGWNTDTFVFGELARRNGIMSDPILLIDIGMLNIRESEK 445
Qy 482 YGFIHLSFORFAAMYIILDEGGGAGPDQVTRLLTEYAFSERSFLALTSR----- 534
Db 446 SYIFLHPSVQEVCAAFYLLKSHVD--HPSQEVKSI-----EKLMAFLKVKVQWIF 496
Qy 535 ---FLGLNEETSHLEKSLWKSVPKSHIKMDLLOWTOSKAOQSDGSTLQOGSLEFFSCLY 591
Db 497 FGSFIFGLHESQKLEAFHGHQLEQIKRQYQCLLETISGNEELQEQIDGKMLFYCLF 556
Qy 592 EIQEEFFIQALSHFOVIVVSNLASK--MEHMYSSFCLEKRCRAOVHLHYGATYSADGD 649
Db 557 EMDDDTFLVEAMCMQI---NFVAKDYSVIVVAHCLKHCFTLKLKLSF-----STQ 605
Qy 650 RARCAGATLLVOLRPERVLLDAYSEHAAALCTNPNLIELSLYRNALSGRGVLLCQ 709
Db 606 NVLSGAQESYM-----ER--LLTAWN-HICSVFIISKQIQLRMDKTNLSGAFSLYN 657
Qy 710 GLRHPNCKLQNLKRCRISSS---ACED--LSAALTANKVLTMDLSGNGVGPFGMMLLC 765
Db 658 NLKHYNTLVN-----VANNVFCEKYLFFELIQNCNLOHNLSTLISHSDVKLL 711
Qy 766 EGLRHPQRLQMTQLRKCOLESACOMASVLGTNPHLVELDLTGNALDELGLRLCOGL 825
Db 712 DVLSQAECEIEELVVAACSISDDCKVFASVLISNKTILKHLNLSNTL-DKGFIASLCKSL 770
Qy 826 RHPVCELRILWLKICBLTAAACDELASTLSVMSQRELDLSNDELGLVLLCEGLRHP 885
Db 771 CHPDCTLEHLVLANCSLNEKWDYLSVLRRNKTLSHLDISSLNDLDEGLKVLCRALTLP 830
Qy 886 TKLOTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLMILAEGLOHPACR 945
Db 831 DSVLSLSLRHCLITISGQDLAEVLRRNNQNLVSLQVSNKKLEDTGVKLLCDIAKHPNCH 890
Qy 946 LOKMLDSCGLTAKACENIYFTLGINQTLTDLTYLTNNALGDTGVRLCKRLSHPGCKLRV 1005
Db 891 LEDLGLAEACELACCEDLASTFTQCKTLWANNLLKNALDYNGLVLCALQOICATVY 950
Qy 1006 LMLFGMLNKMTHSLRALRVTKPYLDI 1033
Db 951 LGLQITDFTETOAFVAEQEKNPCLRI 978

RESULT 10

Q86W28
AC Q86W28 PRELIMINARY; PRT: 1029 AA.
DT 01-JUN-2003 (TREMBLrel. 24, Created)
DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE NALP8.
GN NALP8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA "Schopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
DR EMBL; AY154463; AA018159.1; --
DR InterPro; IPR001611; LRR_RNinh.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh sub.
DR InterPro; IPR007111; NACHT NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR_1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_R1; 7.
DR PROSITE; PS00824; DAPIN; 1.
DR PROSITE; PS00837; NACHT; 1.
SQ SEQUENCE 1029 AA; 117335 MW; 7F8CEC14577303BA CRC64;

Query Match 22.2%; Score 1216.5; DB 4; Length 1029;
Best Local Similarity 33.1%; Pred. No. 1.9e-96;
Matches 325; Conservative 165; Mismatches 391; Indels 101; Gaps 20;
Qy 16 YLBELEAVELKPKKVLGATGELGEGKIP--WGSMEKAGPLEMAQLLITHTFGPEEAURLA 73
Db 44 YMRNVSHLEELQRFKQLL--LTELSTGMPITWQVETASWAEVVLHLLIERPGRADVT 101
Qy 74 LSTFERINRKMWERQEDLVDRDPOETVYRKYRFRMLMEDRNARLGEV--VNLS----- 127
Db 102 SNIFAIMNCDKMCVVVRE-----INAILPTLEPEDLVNGETQVNLGEGESG 148
Qy 128 --HRYTRLLLVKHSNPMQVQQQLLD--TCRGHARTVGHQA---SPIKTIETLEPEEREP 180
Db 149 KIRY-----KSNVNEKFFPDIITWPGNQRFYQGVHREHEYLPCLLLPKRPQ 200
Qy 181 EPPRTVMQGAAGIGKSLAHKVMWADGKLFQGRFDYLFYINCREMNOQATSCSMDL 240
Db 201 RQPKTVAIGAGIGKTLAKVMFEWARKNFVAKWCAFYHCEVQNOT--TDQSFSEL 259
Qy 241 IFSCWEPSPAPIQLIRPERLLFIIDGDELKPSFHDPOGPMWCLWEEKRPTPELLNSL 300
Db 260 IEQWPGSQDLVSKMSKPDQLLLDGFELTSTLIDRLDLSQDWKQLPGSVLLSSL 319
Qy 301 IRKLLPELSLITTRPTALEKHLRLLEHPRHVEILGFSEAEKKEYFYKFFHNAEQGV 360
Db 320 LSKTMLPEATLLIMRFTSWQCKPKLLKCPSLVTLPGFNTMEKIKTFQMYFGHTEGQDV 379
Qy 361 FNYVRDNEPLFMWCFVPLVWVCTCLOOQEGGLLROTSRTTAVMYLLSLMQPCKP 420
Db 380 LSFAMENTILFSMCRVPVVMVCSGLKQOMERGNLITQCPNATSVFVRYISLFPTR- 438
Qy 421 GAPRLQPPNQ---RGLCSLAADGLWNLKILFEEDQRLKHGLDGEDVSFAFLNMFQD 477
Db 439 -AENSFKHQALQELCHLAADSMWRKWLKEDLEAKLDQGTAVTAFGLMSILRITA 497
Qy 478 NCERYYSFIHLSFQBFFAAMYIILDEGGGAGPDQ-----DVTRELLTEYAFSERS 527
Db 498 GEEDHYVTLVTFQFEEFALFYV-----CFQRLKNFHLVSHVNIQLRIASPRGS-KS 550
Qy 528 FLALTSRFLGLNNEETSHLEKSLCWKVSPIHKMDLLOWTOSKAOQSDGSTLQOGSLEFF 587
Db 551 YLSHMGLELFGFLNEACASAVEQSPQCKVSFGNKRKLLKVIPLLHKCDPPSPSGVGPQLF 610
Qy 588 SCLYEIOEEFFIQALSHFOVIVVSNLASKMEHMYSSFCLEKRCRAOVHLHYGATYSADG 647
Db 611 YCLHEIREAFVQSALNDYHKKVVL-RIGNNKEVQVSFAFLKRC---QYLHEVELVTTLNF 666
Qy 648 EDARCSAGATLLVOLRPERVLLDAYSEHAAALCTNPNLIELSLYRNALSGRGVLL 707
Db 667 MNVWKLSSSSHP--GSEAPESNG-LHRWQDLCSVFATNDKLVLTMTNSVLGPFLLKAL 723
Qy 708 CQGLRHPNCKLQNLKRCRISSSACEDLSAALANKVLTMDLSGNGVGPFGMMLLCBG 767
Db 724 AALRHPQCKLQKLLLRVN-STMLNQDLIGVLTGNQHLRYLEIQHVEVESKAVKLCRV 782


```

QY 768 LRHPQCRIQMQLRKQESGACQEMASVLTGNPH----- 802
Db 783 LRSFRCRLQCLREDCIATPRTWIDGNNLQNGHLKTLILRKNLENGAYVLSVAQLE 842
QY 803 -----LVELDLTGNALEDGLRLLCGLRHPVCRLETLMLKICRTATAACDELASTLS 855
Db 843 RLSQSKMLTHLSABNALKDGAHWNALPHLRCPQLRLVLRKCDLTFNCCQDMISALC 902
QY 856 VNQSIRELDLSNELDGLVLLCEGLRHPCKQLQTLRLGICRLGSAACEGLSVVLQANH 915
Db 903 KNTKLSLDLSFNSLKDGVILLCEALKNPCTLIQILENCLFTISICQCAWASMLRKQ 962
QY 916 NLRELDLSFNDLGDWGLWLLAE 937
Db 963 HLRHLDLSKNAIGVYGILTCE 984

RESULT 11
ID Q7TPU9 PRELIMINARY; PRT; 1111 AA.
AC Q7TPU9;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Maternal effect gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ushin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Egg;
RX Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC053384; AAH53384.1; -
SQ SEQUENCE 1111 AA; 125429 MW; ALACBE8E98A546F8 CRC64;

Query Match 21.4%; Score 1168.5; DB 11; Length 1111;
Best Local Similarity 31.5%; Pred. No. 3.3e-92;
Matches 294; Conservative 165; Mismatches 406; Indels 69; Gaps 13;

QY 162 HQASP-1K-1ETLFEDEPERPEPRIVVMGAGIGKSMHAKVMDWADGKLFQGRFDY 219
Db 168 HYDSPENKLSDAFKP-YQKTFQPHYIILHGRFVGKSAARSIVLGNAGKLFQ-KMSF 225
QY 220 LFVINCREMNSQATECSMODLPSFCWPEAPQELIRVPERLLFIIDGDELKPSFHD 279
Db 226 VIFPSVREIKWT-EKSLAQIAKECPDSWDLVTKMSQERLLFVIDGLDDMSVLQHD 284

```

```

QY 280 QGPCLCWEKRPTELLNSLRKLLPELSLLITRPTALEKLRLLEHPRHVEILGFS 339
Db 285 DMTLSRDWKDEQPTIYILMYSLLRKALLPOSELIITRNTGTEKLSKMSVSVPLIYVGLS 344
QY 340 EABEKEYFYKHNAEQAGQVFNVRDNEPLFTMCFVPLVCVWVCTCIQQOLEGGLLRQ 399
Db 345 ASRSQVLVLEINISBESBRIQVHSLIENHQLFDQCQPSVCSINCEALQOKLGRCTL 404
QY 400 TSRTTAVY---MAYLLSLMQPKFQAPRLQPPNORGLCSLAADGLNNQKLFEEQDLRK 456
Db 405 PCQTLTGLYATLVFHLTKRPSQSALSOEQIITLVGLCMAAEGVWTRSVFYVDDDLKN 464
QY 457 HGLGDVSAFLNNI-FKDINCERYYSFTHLSQEFPAAMYLLDSEGGAGDQ--- 512
Db 465 YSLKSEILALFHMNIILQVGNSEQCVFSLSLQDFFAALYYVL---EGLEGWQHFC 521
QY 513 --DVTRLLTEYAFSERSFLALTSRFLGILLNEETRSHLEKSLCWKVSPIHMDLLQWIS 570
Db 522 FIENQRSIMEVKRTDDTFLLGMKRFLLGLMKNKDIILKTLVLFVYFVITVEQKLQHWVSL 581
QY 571 KAQSDGSTLQOQSLEFFSCLYEIOEEPTQOALSHFQVIVVSNIASKMHVSVSCLKRC 630
Db 582 IAAQVNGTSPMDTLDAFYCLFESQDEEFVGALKRFQEWLL-INQKMDLVKSSYCLKHC 640
QY 631 RSAQVHL-----YGATYSADGE----- 648
Db 641 QNLKAIRVIDIRDLSDVNTLPCVVTVOETQCKPLLMWGNFCVLSLRNLKELDGLG 700
QY 649 DRARCSAGAHLLVOLRPE-----RTVLLDAYSEHLAAALCTNPNIELSILYRNAL 699
Db 701 DSILSORAMKILCLERNSQSCRIQKLTFSKAEVVSGLKHLWKLLFSNQNLKYLNGTPE 760
QY 700 SGRGVKLLCOGLRHPNCKLQNLRLKCRISACBDELGAALIANKNLTFMDLSGNGVGP 759
Db 761 KODDMKLAACEALKHPKCSVETLRDSCBLTIIGYEMISTLIISTTRKCLSLAKRNVGK 820
QY 760 GNMILCEGLRHPQCELOMIQRLKCOLESACQEMASVLTGNPHLVELDLTGNALEDGLR 819
Db 821 SMISLGNALSSMCLQKLIIDNCGLTSPASCHLIVSALFNSQNLTHLCSNNSLGTGVQ 880
QY 820 LICOGLRHPVCRTRTLWLKICRLTAAACDELASTLSVNQSLRELDLSNELDGLVLLLC 879
Db 881 QLCQFLRPEALQRLILNHCVNDVDAYGLFAMRLANNTKTLHLSITWNPVGDGAMKLLC 940
QY 880 EGLRHPTCKLQTLRLGICRLGSAACEGLSVVLQANHNIRELDLSFNDLGDWGLWLLAEG 939
Db 941 EALKEPTCYLQLELVDCQLTQNCCEDLACMTITTTKHLKSLDLGNALGDKGVITLCEGL 1000
QY 940 QHPACELQKWLDSCLGTAKACENLYFTGINOTLTDLVLTNNALGDTGVRLLCRSLHP 999
Db 1001 KOSSSLRRKIGUGACKLITSNCCALSLAISNPHLSNLNVKNDFTSGMLKLSAFQCP 1060
QY 1000 GKCLRVLWFLGMDLNKMTSHSLAALRVTKPYLDI 1033
Db 1061 VSNLGIIGLWKQYVYARVRRLQEEVEFKPHVI 1094

RESULT 12
ID Q95LZ7 PRELIMINARY; PRT; 713 AA.
AC Q95LZ7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheciae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:56:27 ; Search time 58 Seconds
(without alignments)
5042.012 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 1035
Sequence: 1 MRLTAGDGLCLSTYLEEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A.Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1035	100.0	1035	5 ADE36417	Human PAA
2	1019	98.5	1035	5 AAO17857	Pyrin dom
3	818	79.0	1061	5 AAO15590	Human PYR
4	661	63.9	865	7 ADC31287	Human nov
5	603	58.3	603	5 ADE36457	Human PAN
6	445	43.0	565	5 ABG97475	Human nuc
7	408	39.4	521	6 ABU99120	Novel hum
8	350	33.8	582	6 ABU99119	Novel hum
9	254	24.5	449	5 ABB77910	Amino aci
10	217	21.0	287	5 AAE21064	Novel hum
11	217	21.0	344	5 AAE21062	Novel hum
12	190	18.4	190	5 ADE36456	Human PAN
13	190	18.4	137	7 ADC32968	Human nov
14	172	16.6	287	5 AAE21063	Novel hum
15	126	12.2	168	4 ABG04570	Novel hum
16	94	9.1	94	4 AAM78275	Human bon
17	81	7.8	159	3 AAB42357	Human ORF
18	75	7.2	75	5 ADE36399	Human PAA
19	56	5.4	56	4 AAM78296	Human bon
20	49	4.7	65	5 AAO15589	Human PYR
21	44	4.3	89	4 AAM83229	Human imm
22	29	2.8	29	5 AAE21075	Novel hum
23	29	2.8	29	5 AAE21074	Novel hum
24	29	2.8	29	5 AAE21072	Novel hum
25	25	2.7	28	5 AAE21071	Novel hum

26	28	2.7	28	5 AAE21068	Novel hum
27	28	2.7	28	5 AAE21069	Novel hum
28	28	2.7	28	5 AAE21070	Novel hum
29	26	2.5	29	5 AAE21073	Novel hum
30	22	2.1	28	5 AAE21067	Novel hum
31	16	1.5	29	5 AAE21076	Novel hum
32	13	1.3	13	5 ABB77911	Protein k
33	13	1.3	13	5 ABB77913	Protein k
34	13	1.3	13	5 ABB77912	Protein k
35	13	1.3	75	5 AAO17871	Novel hum
36	13	1.3	518	5 AAO17871	Novel hum
37	13	1.3	719	5 ADE36452	Human PAA
38	13	1.3	1034	4 AAE07514	Human PYR
39	13	1.3	1034	6 ABU08503	Human PYR
40	13	1.3	1034	6 ABU63315	Human PYR
41	12	1.2	240	4 ABG04098	Novel hum
42	12	1.2	385	5 AAU79523	Human MAT
43	12	1.2	385	6 AAE31746	Human MAT
44	12	1.2	1033	5 ABU65214	Human NOV
45	12	1.2	1143	6 ADA45221	Human MAT

ALIGNMENTS

RESULT 1

ADE36417
ID ADE36417 standard; protein; 1035 AA.

XX ADE36417;

DT 29-JAN-2004 (first entry)

DE Human PAA and nucleotide binding protein PAN6.

XX cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
KW NFkappaB activation inhibitor; PAA domain containing polypeptide;
KW PAA and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NFkappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAA domain family; human; PAN6.

OS Homo sapiens.

PN US2003077699-A1.

XX 24-APR-2003.

PF 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

PR 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

PA (GODZ/) GODZIK A.

PA (CHUZ/) CHU Z.

PA (PAWL/) PAWLOWSKI K.

PA (FIOR/) FIORENTINO L.

PA (ARIZ/) ARIZA M E.

XX (STEH/) STEHLIK C.

PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;
 XX WPI; 2002-471256/50.
 DR N-PSDB; ADE36416.
 XX Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 PS Claim 18; SEQ ID NO 24; 93pp; English.
 XX
 CC The invention describes an isolated PAAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NF-kappaB). A NB-ARC domain polypeptide is useful
 CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (I). (I) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
 CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (I) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6.
 XX
 SQ Sequence 1035 AA;

Query Match 100.0%; Score 1035; DB 5; Length 1035;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1035; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLGATLGEKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGDGLCRSLTYLEELAEVLEKFKLYLGATLGEKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVDPQETRYDYVRKKFRLMEDRNARL 120
 DB 61 ITHFGPEAWRLALSTFERINRKLWERGQREDLVDPQETRYDYVRKKFRLMEDRNARL 120
 QY 121 GECVNLSHRYTRLLVKEHSPNQVQOQLDTGCHARTVGHQASPIKIELFPDDERP 180
 DB 121 GECVNLSHRYTRLLVKEHSPNQVQOQLDTGCHARTVGHQASPIKIELFPDDERP 180
 QY 181 EPPRTVMQAGAGKSLAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSQDQL 240
 DB 181 EPPRTVMQAGAGKSLAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSQDQL 240
 QY 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQPCWLCWEERKPTLLNSL 300
 DB 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPSPHDPQPCWLCWEERKPTLLNSL 300
 QY 301 IRKLLPELSLITTRTALEKHLRLEHPRHVEILGFSEARKEYFYKYPHNAEQGV 360
 DB 301 IRKLLPELSLITTRTALEKHLRLEHPRHVEILGFSEARKEYFYKYPHNAEQGV 360
 QY 361 FNYVRDNEPLFTMCFVPLVCVWCTCQQQLGEGGLLRQTSRTTAVYMLYLLSIMQPKP 420
 DB 361 FNYVRDNEPLFTMCFVPLVCVWCTCQQQLGEGGLLRQTSRTTAVYMLYLLSIMQPKP 420
 QY 421 GAPRLQPPPNORGLCSLAADGLMNQKILFERQDLRKXGLDGEDVSAFLNMNIFOKDNCE 480

DB 421 GAPRLQPPPNORGLCSLAADGLMNQKILFERQDLRKXGLDGEDVSAFLNMNIFOKDNCE 480
 QY 481 RYYSFTHLSQBFPAAMYIILDEGEGAGAPDQVTRLLITAYASERSFIALTSRFLGILL 540
 DB 481 RYYSFTHLSQBFPAAMYIILDEGEGAGAPDQVTRLLITAYASERSFIALTSRFLGILL 540
 QY 541 NEETRSHLEKSLCWKSPHIMKMDLLOWIQSKAOSDGTSTLOQGSLEFFSCYIETQEEBFIQ 600
 DB 541 NEETRSHLEKSLCWKSPHIMKMDLLOWIQSKAOSDGTSTLOQGSLEFFSCYIETQEEBFIQ 600
 QY 601 QALSHFOVIVVSNIAKMEHMSVFCFKRCSAQVLHLYGATYSADGEDRARSAGAHTL 660
 DB 601 QALSHFOVIVVSNIAKMEHMSVFCFKRCSAQVLHLYGATYSADGEDRARSAGAHTL 660
 QY 661 LVQLRPERTVLLDAYSEHLLAAALCTNPNIETLSYRNALSGRGVLLCOGLRHPNCKLON 720
 DB 661 LVQLRPERTVLLDAYSEHLLAAALCTNPNIETLSYRNALSGRGVLLCOGLRHPNCKLON 720
 QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFFPGMWMLICEGLRHPQCRLOMIQL 780
 DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNVGFFPGMWMLICEGLRHPQCRLOMIQL 780
 QY 781 RKQLESAGACQEMASVLTGNPHLVDELDTGNALEDLGLRLCCOGLRHPVCLRTLWLKIC 840
 DB 781 RKQLESAGACQEMASVLTGNPHLVDELDTGNALEDLGLRLCCOGLRHPVCLRTLWLKIC 840
 QY 841 RLTAACDELASTLSVYNQSLRELDLSNELGDLGVLLICEGLRHPVCLRTLWLKIC 900
 DB 841 RLTAACDELASTLSVYNQSLRELDLSNELGDLGVLLICEGLRHPVCLRTLWLKIC 900
 QY 901 SAACEGLSVVLQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKLDSCGLTAKA 960
 DB 901 SAACEGLSVVLQANHNRLDLSFNDLGDWGLWLLAEGLQHPACRLQKLDSCGLTAKA 960
 QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPGCKLRVLWFGMDLNKMTSHR 1020
 DB 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPGCKLRVLWFGMDLNKMTSHR 1020
 QY 1021 LAALRVTKPYLDIGC 1035
 DB 1021 LAALRVTKPYLDIGC 1035
 RESULT 2
 ID AAO17857 standard; protein; 1099 AA.
 AC AAO17857;
 XX 20-AUG-2002 (first entry)
 XX Pyrin domain containing protein NALP3/PY5-hs.
 XX Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
 XX antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
 XX neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
 XX nephrotropic; osteoparactic; nootropic; intracellular signal transduction;
 XX inflammation; Alzheimer's disease; infection; psoriasis; asthma;
 XX arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
 XX osteoarthritis; glomerulonephritis.
 XX Unidentified.
 XX W0200240668-A2.
 XX 23-MAY-2002.
 XX 30-OCT-2001; 2001WO-EP012545.
 XX 15-NOV-2000; 2000DE-01056687.
 XX 30-NOV-2000; 2000DE-01059595.

PA (APOT-) APOTECH RES & DEV LTD.
 XX Tschoop J, Marton F;
 DR WPI; 2002-427093/45.
 DR N-PSDB; AAL47129.
 XX
 PT New DNA encoding protein with pyrin domain, useful for treating diseases
 PT involving impaired signal transduction, particularly inflammation, also
 PT proteins and antibodies.
 XX
 PS Claim 5; Fig 1; 116pp; German.
 XX
 CC The present invention relates the DNA and their encoded proteins, where
 CC the proteins contain at least one PYD (pyrin) domain. These can be used
 CC to treat diseases associated with impaired intracellular signal
 CC transduction, particularly inflammation such as psoriasis,
 CC arteriosclerosis, bacterial or viral infections (particularly meningitis
 CC and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma,
 CC sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's
 CC and Parkinson's diseases. The present sequence is a protein of the
 CC invention
 XX
 SQ Sequence 1099 AA;

Query Match 98.5%; Score 1019; DB 5; Length 1099;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1019; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLRTAGDGLCRISTYLEEAEVLEAEVKFKLYLTATLGEKIPWGSMEKAGPLEMAQLL 60
 DB 1 MLRTAGDGLCRISTYLEEAEVLEAEVKFKLYLTATLGEKIPWGSMEKAGPLEMAQLL 60
 QY 61 ITHFGPEAWRLALSTFERINRKNLWEGQREDLVRDPQETRYDYVRKPRMEDRNARL 120
 DB 61 ITHFGPEAWRLALSTFERINRKNLWEGQREDLVRDPQETRYDYVRKPRMEDRNARL 120
 QY 121 GECVNLSHRYTRILLVKEHNPQVQOQLDTRGHARTVGHQASPIKIETLPEPDEEP 180
 DB 121 GECVNLSHRYTRILLVKEHNPQVQOQLDTRGHARTVGHQASPIKIETLPEPDEEP 180
 QY 181 EPRTVMQGAAGIGKSMIAHKVMDWADKLFQGRFDYLFYINCREMNQSAECSQMDL 240
 DB 181 EPRTVMQGAAGIGKSMIAHKVMDWADKLFQGRFDYLFYINCREMNQSAECSQMDL 240
 QY 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLXPSFHPDQGPWCLCWEKRPTELLINSL 300
 DB 241 IFSCWPEPSAPLQELIRVPERLLFIIDGFDLXPSFHPDQGPWCLCWEKRPTELLINSL 300
 QY 301 IRKKLPELSLITTRPTALEKHLRLLEHPRHVEILGFSEAEKVEYKYFHNAGQOV 360
 DB 301 IRKKLPELSLITTRPTALEKHLRLLEHPRHVEILGFSEAEKVEYKYFHNAGQOV 360
 QY 361 FNYVRNEPLFTMCFVPLVCWVCTCIQQOEGGLLRQTSRTTAVYMLYLLSLMQPKP 420
 DB 361 FNYVRNEPLFTMCFVPLVCWVCTCIQQOEGGLLRQTSRTTAVYMLYLLSLMQPKP 420
 QY 421 GAPRLQPPNQRGLSLAADGLWNQKILFPEQDLRKHLGDGSDVSAFLNNIFQKQINCE 480
 DB 421 GAPRLQPPNQRGLSLAADGLWNQKILFPEQDLRKHLGDGSDVSAFLNNIFQKQINCE 480
 QY 481 RYYSFTHLSFOERFAAMYVILDEGCGAGPDQDVTLLTTEYAFSEFSFLATLSRELGLL 540
 DB 481 RYYSFTHLSFOERFAAMYVILDEGCGAGPDQDVTLLTTEYAFSEFSFLATLSRELGLL 540
 QY 541 NEETRSHLEKSLCWKVSPIHKMOLLQWISKAQSDGSTLQQSGLEFFSCLYEQEEFFIQ 600
 DB 541 NEETRSHLEKSLCWKVSPIHKMOLLQWISKAQSDGSTLQQSGLEFFSCLYEQEEFFIQ 600
 QY 601 QALSHFOVIVVSNIAKMEHNVSSFCILKRCRQAQVHLHYGATYSADGEDRARSAGAHTL 660
 DB 601 QALSHFOVIVVSNIAKMEHNVSSFCILKRCRQAQVHLHYGATYSADGEDRARSAGAHTL 660

QY 661 LVOLRPERTVLLDAYSEHLAAALCTNPNIELSLYRNALSGRGVKKLLCOGLRHPNCKLQ 720
 DB 661 LVOLRPERTVLLDAYSEHLAAALCTNPNIELSLYRNALSGRGVKKLLCOGLRHPNCKLQ 720
 QY 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLRHPQCRLOMIQL 780
 DB 721 LRLKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLCEGLRHPQCRLOMIQL 780
 QY 781 RKQLESAGCOEMASVLTGNPHLVDELDTGNALEDLGLRLCOGLRHPVCRILTLWLKIC 840
 DB 781 RKQLESAGCOEMASVLTGNPHLVDELDTGNALEDLGLRLCOGLRHPVCRILTLWLKIC 840
 QY 841 RLTAACDELASTLSVNSQSLRELDLSNELGDGLVLLCEGLRHPCKLQTLRLGICRLG 900
 DB 841 RLTAACDELASTLSVNSQSLRELDLSNELGDGLVLLCEGLRHPCKLQTLRLGICRLG 900
 QY 901 SAACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGGLTAKA 960
 DB 901 SAACEGLSVVLOAHNHLRELDLSFNDLGDWGLWLLAEGLOHPACRLQKWLDSGGLTAKA 960
 QY 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPCCKLRVLWLFQMDLNKMTS 1019
 DB 961 CENLYFTLGINQTLTDLYLTNNALGDTGVRLCKRLSHPCCKLRVLWLFQMDLNKMTS 1019

RESULT 3
 AAO15590
 ID AAO15590 standard; protein; 1061 AA.
 XX
 AC AAO15590;
 DT 31-OCT-2002 (first entry)
 XX
 DE Human PYRIN-8 protein #2.
 XX
 KW Human; gene therapy; PYRIN; stress-related response; apoptotic response;
 KW inflammatory response; inflammatory disorder; immune system disorder;
 KW Crohn's disease; multiple sclerosis; cancer; leukaemia;
 KW autoimmune disorder; arthritis; neurological disease;
 KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
 KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
 KW transcription profiling; PYRIN-8.
 XX
 OS Homo sapiens.
 XX
 PN WO200261049-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002967.
 XX
 PR 31-JAN-2001; 2001US-0265231P.
 PR 10-SEP-2001; 2001US-0318645P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PA (AMHP) WYETH.
 XX
 PI Bertin J, Wang W, Blatcher M;
 XX
 DR WPI; 2002-627477/67.
 DR N-PSDB; AAL44363.
 XX
 PT New PYRIN polypeptides and nucleic acids useful for modulating and
 PT diagnosing stress-related, apoptotic and inflammatory responses, or for
 PT treating inflammatory and immune system disorders, cancers, or
 PT neurological diseases.
 XX
 PS Claim 8; Fig 8; 167pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of human
 CC PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
 CC useful for modulating and diagnosing stress-related, apoptotic and
 CC inflammatory responses. The PYRIN protein and DNA sequences are useful

CC for treating: inflammatory disorders and immune system disorders (e.g.
 CC Crohn's disease, reactive arthritis; multiple sclerosis, contact
 CC dermatitis, psoriasis, graft rejection, allergies, viral infections and
 CC bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
 CC (e.g. systemic lupus erythematosus and arthritis); and neurological
 CC diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
 CC protein and DNA sequences may also be used in screening assays, detection
 CC assays (e.g. chromosomal mapping, tissue typing or forensic biology),
 CC predictive medicine (e.g. diagnostic assays, clinical trials and
 CC pharmacogenomics) and transcription profiling. The present amino acid
 CC sequence represents a human PYRIN-8 protein
 XX
 SQ Sequence 1061 AA;

Query Match 79.0%; Score 818; DB 5; Length 1061;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 938; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 97 DPQTYRDYVRKFRMLMEDRNARLGCYNLSHRVTRLLLVKHEHNPVQVQQLDTCGRH 156
 Db 124 DPQTYRDYVRKFRMLMEDRNARLGCYNLSHRVTRLLLVKHEHNPVQVQQLDTCGRH 183
 QY 157 ARTVGHQASPIKIEITLPEDEERPEPRTPVVMQGAAGIGKSMIAHKVMDWADKLFQGR 216
 Db 184 ARTVGHQASPIKIEITLPEDEERPEPRTPVVMQGAAGIGKSMIAHKVMDWADKLFQGR 243
 QY 217 FDYLFYINCRMNQASQCSMODLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPFS 276
 Db 244 FDYLFYINCRMNQASQCSMODLIFSCWPEPSAPLQELIRVPERLLFIIDGFDLKPFS 303
 QY 277 HDQPGMCLCWEERKPELLNLSLRKLLPELSLITRPTALEKLRHLEHPRHVEIL 336
 Db 304 HDQPGMCLCWEERKPELLNLSLRKLLPELSLITRPTALEKLRHLEHPRHVEIL 363
 QY 337 GFSEARKEFYKYFHNAQAGQVFNVRDNEPLFTMCVFPLVCWVYTCLOQLQEGGL 396
 Db 364 GFSEARKEFYKYFHNAQAGQVFNVRDNEPLFTMCVFPLVCWVYTCLOQLQEGGL 423
 QY 397 LRQTSRTTAVYMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEEQDLRK 456
 Db 424 LRQTSRTTAVYMLYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFEEQDLRK 483
 QY 457 HGLDGEDVSAFLNWNIFQKIDNCERYYSFHLSEQEFPAAMYIILDEGEGAGPDQDVT 516
 Db 484 HGLDGEDVSAFLNWNIFQKIDNCERYYSFHLSEQEFPAAMYIILDEGEGAGPDQDVT 543
 QY 517 LLTEYAFSRSFALTGRFLGLNEETRSLKSLCKWVSPHIMDLOWIQSAQSDG 576
 Db 544 LLTEYAFSRSFALTGRFLGLNEETRSLKSLCKWVSPHIMDLOWIQSAQSDG 603
 QY 577 STLOQGSLEFPSCLEYEQEEPIQALSHFQVIVVSNIAKWEHNVSSFCILKRCESAOVL 636
 Db 604 STLOQGSLEFPSCLEYEQEEPIQALSHFQVIVVSNIAKWEHNVSSFCILKRCESAOVL 663
 QY 637 HLYGATYSADGEDRARSAGATHLLVQLRPRTRVLLDAYSEHLAAALCTNPNLIELSLYR 696
 Db 664 HLYGATYSADGEDRARSAGATHLLVQL-PRTRVLLDAYSEHLAAALCTNPNLIELSLYR 722
 QY 697 NALGSRGVKLLCOGLRHNPCKLQNLRLKRCISSACEDLSAALIANKNLTMDSNGV 756
 Db 723 NALGSRGVKLLCOGLRHNPCKLQNLRLKRCISSACEDLSAALIANKNLTMDSNGV 782
 QY 757 GFPGMMLLCEGLRHQPCKLQNLRLKRCISSACEDLSAALIANKNLTMDSNGV 816
 Db 783 GFPGMMLLCEGLRHQPCKLQNLRLKRCISSACEDLSAALIANKNLTMDSNGV 842
 QY 817 GLRLCCQGLRHVPCKRLTLWLKICRLTAACDELASTLSVNQSLRELDLSNELGDLVL 876
 Db 843 GLRLCCQGLRHVPCKRLTLWLKICRLTAACDELASTLSVNQSLRELDLSNELGDLVL 902
 QY 877 LLCEGLRHPTCKLQTLRLGICELGSAACEGIVSVTLQANHNRLRELDLSFNDLGDWGLWLLA 936
 Db 903 LLCEGLRHPTCKLQTLRLGICELGSAACEGIVSVTLQANHNRLRELDLSFNDLGDWGLWLLA 962

QY 937 EGLQHPACRLQKWLDSGGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLICKRL 996
 Db 963 EGLQHPACRLQKWLDSGGLTAKACENLYFTLGINQTLTDLTLTNALGDTGVRLICKRL 1022
 QY 997 SHFGCKLRVLWLFQGMDLNKNMTHSRALALRVTKPYLDIGC 1035
 Db 1023 SHFGCKLRVLWLFQGMDLNKNMTHSRALALRVTKPYLDIGC 1061
 RESULT 4
 ADC31287
 ID ADC31287 standard; protein; 865 AA.
 XX
 AC ADC31287;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 DE Human novel polypeptide sequence, SEQ ID NO:1369.
 XX
 KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
 KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 19.
 XX
 OS Homo sapiens.
 XX
 WO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PP 24-SEP-2002; 2002WO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC30316.
 PT
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 FS Claim 20; SEQ ID NO 1369; 1185pp; English.
 CC
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the

CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 XX Sequence 865 AA;
 Query Match 63.9%; Score 661; DB 7; Length 865;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 781; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	113	MEDNRALGECVNLSHRYTRLLLVKEHSPMQVQQQLLTGRGHARTVGHQASPIKIETL	172
Db	1	MEDNRALGECVNLSHRYTRLLLVKEHSPMQVQQQLLTGRGHARTVGHQASPIKIETL	60
QY	173	FPDDEPPEPTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFYLYINCREMNQSA	232
Db	61	FPDDEPPEPTVVMQGAAGIGKSMIAHKVMDWADGKLFQGRFYLYINCREMNQSA	120
QY	233	TSCSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPDQPGPWCCLWEKEKP	292
Db	121	TSCSMODLIFSCWPEPSAPLOELIRVPERLLFIIDGFDLKPSPDQPGPWCCLWEKEKP	180
QY	293	TELLNSLRKLLPELSLLITRPTALEKHLRLEHPRHVEITLGFSEAEKKEYFYKYFH	352
Db	181	TELLNSLRKLLPELSLLITRPTALEKHLRLEHPRHVEITLGFSEAEKKEYFYKYFH	240
QY	353	NAEQAGQVNYVRDNEPLFTMCPVPLVWVCTCLOQLEGGGLLAQTSTTTAVYMLYL	412
Db	241	NAEQAGQVNYVRDNEPLFTMCPVPLVWVCTCLOQLEGGGLLAQTSTTTAVYMLYL	300
QY	413	LSLMQPKGAPRIOPPNEGRLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSFAFLNMI	472
Db	301	LSLMQPKGAPRIOPPNEGRLCSLAADGLWNQKILFEEQDLRKHGLDGEDVSFAFLNMI	360
QY	473	FOKDINCERYYSFIHLSFOEFFAAMYIILDEGGGAGPDQDVRLLTTEYAFSRSFLALT	532
Db	361	FOKDINCERYYSFIHLSFOEFFAAMYIILDEGGGAGPDQDVRLLTTEYAFSRSFLALT	420
QY	533	SRFLGGLNNEETRSHLEKSLCWKVSPIHKMDLLOWTOSKAQSDGSTLQOQSLEFFSCLYE	592
Db	421	SRFLGGLNNEETRSHLEKSLCWKVSPIHKMDLLOWTOSKAQSDGSTLQOQSLEFFSCLYE	480
QY	593	IQBEETIQALSHFQVIVVSNASKMEHVMVSSPCLAKRCSAQLVHLGYATYSADGEDRAR	652
Db	481	IQBEETIQALSHFQVIVVSNASKMEHVMVSSPCLAKRCSAQLVHLGYATYSADGEDRAR	540
QY	653	CSAGATLLVQLRPETVILLDAYSEHLAAALCTNPMLIELSLYRNALGSRGVKLLCOGLR	712
Db	541	CSAGATLLVQL-PERTVILLDAYSEHLAAALCTNPMLIELSLYRNALGSRGVKLLCOGLR	599
QY	713	HPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTMDLSGNGVGFPGMWMLCEGRHPQ	772
Db	600	HPNCKLQNLRLKRCRISSSACEDLSAALIANKNLTMDLSGNGVGFPGMWMLCEGRHPQ	659
QY	773	CRLOMTQLRKQCESGACQEMASVLTGNPHLVELDTGNALEDGLRLLCQGLRHPVCRL	832
Db	660	CRLOMTQLRKQCESGACQEMASVLTGNPHLVELDTGNALEDGLRLLCQGLRHPVCRL	719
QY	833	RTLWLKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGRHPTCKLQTL	892
Db	720	RTLWLKICRLTAACDELASTLSVNSQSLRELDLSINELGDLGVLLICEGRHPTCKLQTL	779

QY	893	RL 894
Db	780	RL 781

RESULT 5
 ADE36457
 ID ADE36457 standard; protein; 603 AA.
 XX
 AC ADE36457;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human PAN6 leucine-rich-repeat domain (LRR) seq id 64.
 XX
 KW cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
 KW antiallergic; antitumor; dermatological; cerebroprotective; cardiact;
 KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
 KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
 KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
 KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
 KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
 KW apoptosis; NFkappaB induction; cytokine processing;
 KW cytokine receptor signaling caspase-mediated proteolysis;
 KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
 KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
 KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
 KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
 KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
 KW arthritis; lupus; schroen's syndrome; Crohn's disease;
 KW ulcerative colitis; graft versus host disease; stroke; heart failure;
 KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
 KW cancer therapy; PAAD domain family; human; PAN6; leucine-rich-repeat;
 LRR.
 XX
 OS Homo sapiens.
 XX
 PN US2003077699-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 25-SEP-2001; 2001US-00965621.
 XX
 PR 26-SEP-2000; 2000US-00671760.
 PR 26-SEP-2000; 2000US-0367367E.
 XX
 PA (REED/J) REED J C.
 PA (GODZ/) GODZIK A.
 PA (CHUZ/) CHU Z.
 PA (PAWL/) PAWLOWSKI K.
 PA (FIOR/) FIORENTINO L.
 PA (ARIZ/) ARIZA M E.
 PA (STEH/) STEHLIK C.
 XX
 PI Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
 PI Stehlik C;
 XX
 DR WPI; 2002-471256/50.
 XX
 PT Novel isolated PAAD domain containing polypeptide useful for inducing
 PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
 PT therapy for treating cancer.
 XX
 PS Claim 24; SEQ ID NO 64; 93pp; English.
 XX
 CC The invention describes an isolated PAAD domain containing polypeptide
 CC (I) comprising 80% identity to the amino acid sequence of PAAD and
 CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
 CC -like protein containing a caspase recruitment domain (ASC)-2 fully
 CC defined in specification, where (I) is biologically active. (I) is useful
 CC for identifying a (I)-associated polypeptide, an agent altering that
 CC association and agents that modulate PAAD domain mediated inhibition of
 CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful

CC for identifying an agent that modulates the activity of the NB-ARC domain
 CC of (1). (1) or its functional fragments is useful in altering cellular or
 CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
 CC processing, cytokine receptor signaling, caspase-mediated proteolysis or c
 CC -Jun N-terminal kinase activation, thus having modulating effect on cell
 CC life and death (apoptosis) inflammation, cell adhesion or other cellular
 CC or biochemical processes. (1) is useful for treating cancer pathologies,
 CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
 CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
 CC proliferation in arteries following balloon angioplasty (restenosis),
 CC leukemia, lymphomas; inflammatory diseases such as allergies, arthritis,
 CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
 CC versus host disease, stroke, heart failure, neurodegenerative diseases
 CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
 CC infection (HIV). (1) is useful for diagnosing cancer or monitoring cancer
 CC therapy. This is the amino acid sequence of a human PAAD and nucleotide
 CC binding protein PAN6 leucine-rich-repeat domain (LLR).
 CC
 SQ Sequence 603 AA;

Query Match 58.3%; Score 603; DB 5; Length 603;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 603; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 429 PNORGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNNMIFQKINCERYYSIHL 488
 DB 1 PNORGLCSLAADGLWNOKILFEEODLRKHGLDGEDVSFAFLNNMIFQKINCERYYSIHL 60
 QY 489 SPOEFFAAMYIILDEGGAGGPDQDVTLLTEYAFSERSFLATSRFLGLNNEETSHL 548
 DB 61 SPOEFFAAMYIILDEGGAGGPDQDVTLLTEYAFSERSFLATSRFLGLNNEETSHL 120
 QY 549 EKSICWKVSPHIXMDLLOWTOSKASQDSTLQOGLSEFFSCLYEIOEEFTQOALSHPQV 608
 DB 121 EKSICWKVSPHIXMDLLOWTOSKASQDSTLQOGLSEFFSCLYEIOEEFTQOALSHPQV 180
 QY 609 IVVSNIAKMEHVMVSSFLKRCRQAQVHLGYATYSADGEDRACRSAGAHVLLVQLRPER 668
 DB 181 IVVSNIAKMEHVMVSSFLKRCRQAQVHLGYATYSADGEDRACRSAGAHVLLVQLRPER 240
 QY 669 TVLLDAYSEHLAALCTNPMLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRI 728
 DB 241 TVLLDAYSEHLAALCTNPMLIELSLYRNALGSRGVKLLCQGLRHPNCKLQNLRLKRCRI 300
 QY 729 SSSACEDLSAALTANKNLTRMDLSGNGVGPWMLLCEGLRHPNCKLQNLRLKRCRI 788
 DB 301 SSSACEDLSAALTANKNLTRMDLSGNGVGPWMLLCEGLRHPNCKLQNLRLKRCRI 360
 QY 789 ACQEMASVLGTNPHVLDLTGNALEDLGLRLCQGLRHPNCKLQNLRLKRCRI 848
 DB 361 ACQEMASVLGTNPHVLDLTGNALEDLGLRLCQGLRHPNCKLQNLRLKRCRI 420
 QY 849 ELASTLSVNSQSLRELDLSINELGDLGVLLICEGLRHPNCKLQNLRLKRCRI 908
 DB 421 ELASTLSVNSQSLRELDLSINELGDLGVLLICEGLRHPNCKLQNLRLKRCRI 480
 QY 909 VVLQANHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKMLDSCGLTAKACENLYFTL 968
 DB 481 VVLQANHLRELDLSFNDLGDWGLWLAEGLOHPACRLQKMLDSCGLTAKACENLYFTL 540
 QY 969 GINQTLTDLVLTNNALGDTGVRLLCCKRLSHPGCKLRLVLMFGMDLNKMTSHRLAALRVTK 1028
 DB 541 GINQTLTDLVLTNNALGDTGVRLLCCKRLSHPGCKLRLVLMFGMDLNKMTSHRLAALRVTK 600
 QY 1029 PYL 1031
 DB 601 PYL 603

RESULT 6
 ABG97475
 ID ABG97475 standard; protein; 565 AA.
 XX

AC ABG97475;
 XX
 DT 16-DEC-2002 (first entry)
 XX
 DE Human nucleic acid associated protein, NAAP9, from, INCYTE no.429930CD1.
 XX
 KW Human; nucleic acid associated protein; NAAP; cancer;
 KW cell proliferative disease; cancer; atherosclerosis; hepatitis;
 KW neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
 KW epilepsy; developmental disorder; renal tubular acidosis; anaemia;
 KW glaucoma; hypothyroidism; autoimmune disorder; AIDS;
 KW inflammatory disorder; acquired immunodeficiency syndrome; allergy;
 KW atopic dermatitis; arthritis; bacterial infection; viral infection;
 KW parasitic infection; protozoal infection; fungal infection.
 XX
 OS Homo sapiens.
 XX
 PN WO200272630-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 07-FEB-2002; 2002WO-US003844.
 XX
 PR 09-FEB-2001; 2001US-0268118P.
 PR 21-FEB-2001; 2001US-0270963P.
 PR 22-FEB-2001; 2001US-0270858P.
 PR 23-FEB-2001; 2001US-0271194P.
 PR 07-MAR-2001; 2001US-0274071P.
 PR 12-APR-2001; 2001US-0283496P.
 PR 09-NOV-2001; 2001US-0344650P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
 PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
 PI Gandhi AR, Ding L, Yue H, Gietzen KU, Wallia NK, Thangavelu K;
 PI Elliott VS, Marquis JP;
 XX
 DR WPI: 2002-723320/78.
 DR N-PSDB; ABS78719.
 XX
 FT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
 FT treating and preventing diseases or conditions associated with the
 FT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
 FT infections.
 XX
 PS Claim 1; Page 147-149; 162pp; English.
 XX
 CC The invention relates to an isolated polypeptide comprising one of 10
 CC human nucleic acid associated protein (NAAP1-10), or a biologically
 CC active or immunogenic fragment of the polypeptide, and their encoding
 CC nucleic acid. Also included are a recombinant polynucleotide comprising a
 CC promoter sequence operably linked to the polynucleotide, a cell
 CC transformed with the recombinant polynucleotide, a transgenic organism
 CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
 CC screening for a compound that is effective as an ant/agonist or modulator
 CC of NAAP, generating an expression profile of a sample containing the
 CC polynucleotides and an array comprising different nucleotide molecules
 CC affixed on a solid substrate, nucleotide molecule specifically hybridisable with
 CC oligonucleotide or polynucleotide sequence specifically hybridisable with
 CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
 CC The polypeptides and polynucleotides are useful in diagnosing, treating
 CC and preventing diseases or conditions associated with the decreased
 CC expression or overexpression of NAAP, such as cell proliferative diseases
 CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
 CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
 CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
 CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired
 CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
 CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
 CC other diseases and disorders listed in the specification. These are also
 CC useful in assessing the effects of exogenous compounds on the expression
 CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its

CC fragments are useful in screening compounds for effectiveness as agonist
 CC or antagonist of the polypeptides, or in altering the expression of the
 CC target polynucleotide and compounds that specifically bind to or modulate
 CC the activity of the polypeptide. The microarray is useful in monitoring
 CC or measuring protein-protein interactions, drug-target interactions, and
 CC gene expression profiles. The present sequence represents an NAAP protein
 XX
 SQ Sequence 565 AA;

Query Match 43.0%; Score 445; DB 5; Length 565;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 565; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 470 MNIFQKINCERYYSFIHLSFQFFRAAMYIILDEGGGAGPDQVTRLLTEYAFSRSYL 529
 Db 1 MNIFQKINCERYYSFIHLSFQFFRAAMYIILDEGGGAGPDQVTRLLTEYAFSRSYL 60
 QY 530 ALTSRFLGLLNEETRSHEKSLCKWVSPHIKMDLQWIOSKAQSDGSTLQOQSLEFFSC 589
 Db 61 ALTSRFLGLLNEETRSHEKSLCKWVSPHIKMDLQWIOSKAQSDGSTLQOQSLEFFSC 120
 QY 590 LYEIOEEFIQALSHFQVIVVSNIAKMEHMYVSFCLKRCRQAQVHLHYGATYSADGED 649
 Db 121 LYEIOEEFIQALSHFQVIVVSNIAKMEHMYVSFCLKRCRQAQVHLHYGATYSADGED 180
 QY 650 RARCSAGATLLVQLPERTVLLDAYSEHLAALCTNPNIJELSLYRNALGSGVKKLQ 709
 Db 181 RARCSAGATLLVQLPERTVLLDAYSEHLAALCTNPNIJELSLYRNALGSGVKKLQ 239
 QY 710 GLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPMMLCEGLR 769
 Db 240 GLRHPNCKLQNLKRCRISSSACEDLSAALIANKNLTRMDLSGNGVGFPMMLCEGLR 299
 QY 770 HPQCRLOMTQLKRCQESGACQEMASVLTGNPHLVELDTGNALDGLRLLCQGLRHPV 829
 Db 300 HPQCRLOMTQLKRCQESGACQEMASVLTGNPHLVELDTGNALDGLRLLCQGLRHPV 359
 QY 830 CRLRTLWLKICRLTAACDELASTLSVNSQSLRELDLSNELGDLGVLLICEGRHPCTCKL 889
 Db 360 CRLRTLWLKICRLTAACDELASTLSVNSQSLRELDLSNELGDLGVLLICEGRHPCTCKL 419
 QY 890 QTLRLGICRLGSAACGLSVLQANLNRLDLDSFNDLGDWGLWLLAEGLOHPACRLQKL 949
 Db 420 QTLRLGICRLGSAACGLSVLQANLNRLDLDSFNDLGDWGLWLLAEGLOHPACRLQKL 479
 QY 950 WLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRSLHPGCKLRVLWLF 1009
 Db 480 WLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRSLHPGCKLRVLWLF 539
 QY 1010 GMDLNKWTSHRLAALRVTKPYLDIGC 1035
 Db 540 GMDLNKWTSHRLAALRVTKPYLDIGC 565

RESULT 7
 ID ABU99120 standard; protein; 521 AA.
 XX AC ABU99120;
 XX DT 01-AUG-2003 (first entry)
 XX DE Novel human GPCR related protein NOV2b.
 XX Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytosolic; cardiac; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antilasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; AIDS; obesity; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;

KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.
 OS Homo sapiens.
 PN WO200299116-A2.
 XX 12-DEC-2002.
 PD 04-JUN-2002; 2002WO-US017428.
 XX 04-JUN-2001; 2001US-0295607P.
 PR 04-JUN-2001; 2001US-0295661P.
 PR 06-JUN-2001; 2001US-0296404P.
 PR 06-JUN-2001; 2001US-0296418P.
 PR 14-JUN-2001; 2001US-0298285P.
 PR 15-JUN-2001; 2001US-0298556P.
 PR 21-JUN-2001; 2001US-0299949P.
 PR 26-JUN-2001; 2001US-0300883P.
 PR 28-JUN-2001; 2001US-0301550P.
 PR 13-AUG-2001; 2001US-0311972P.
 PR 27-AUG-2001; 2001US-0315071P.
 PR 29-AUG-2001; 2001US-0315660P.
 PR 14-SEP-2001; 2001US-0322293P.
 PR 17-SEP-2001; 2001US-0322706P.
 PR 14-DEC-2001; 2001US-0341186P.
 PR 28-FEB-2002; 2002US-0361189P.
 PR 12-MAR-2002; 2002US-0363673P.
 PR 12-MAR-2002; 2002US-0363676P.
 PR 03-JUN-2002; 2002US-00363676.
 XX (CURA-) CURAGEN CORP.
 PA Anderson DW, Baumgartner JC, Boldog FL, Casman SU, Edinger SR;
 PI Ganggalli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;
 PI Pena CE, Rastelli L, Shinkets RA, Stone DJ, Spytek KA, Vernet CAM;
 PI Voss EZ, Zerhusen BD;
 XX MPI; 2003-140627/13.
 DR N-PSDB; ACD03624.
 XX New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy, or
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX Claim 1; Page 100; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods. The nucleic acids are further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The polypeptides are also useful as
 CC vaccines. This is the amino acid sequence of a novel human G-protein
 CC coupled receptor related protein NOV
 XX Sequence 521 AA;

Query Match		39.4%; Score 408; DB 6; Length 521;
Best Local Similarity		100.0%; Pred. No. 0;
Matches 408; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	373	MCVPLVAVVCTCLOQLEGGGLLQTSRTTAVVYLLSLMQPKGAPRLQPPNQ 432
DB	1	MCVPLVAVVCTCLOQLEGGGLLQTSRTTAVVYLLSLMQPKGAPRLQPPNQ 60
QY	433	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE 492
DB	61	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE 120
QY	493	FFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFLATSRFLGLNETHSHLEKSL 552
DB	121	FFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFLATSRFLGLNETHSHLEKSL 180
QY	553	CKVSPHIKMDLLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHFQIVVS 612
DB	181	CKVSPHIKMDLLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHFQIVVS 240
RESULT 8		
ID	ABU99119	standard; protein; 582 AA.
XX	AC	ABU99119;
XX	AC	ABU99119;
DT	01-AUG-2003	(first entry)
XX	DE	Novel human GPCR related protein NOV2a.
KW	Human; G-protein coupled receptor related protein; GPCR related protein;	
KW	NOV; cytosolic; cardiant; antiarteriosclerotic; antidiabetic;	
KW	immunomodulator; anti-HIV; anorectic; antidiabetic; haemostatic;	
KW	antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;	
KW	NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;	
KW	diabetes; immune disorder; AIDS; obesity; asthma;	
KW	haematopoietic disorder; Parkinson's disease; Alzheimer's disease;	
KW	infection; multiple sclerosis; cancer-associated cachexia;	
KW	wasting disorder; chronic disease; neurogenesis; cell differentiation;	
KW	cell proliferation; haematopoiesis; wound healing; angiogenesis;	
KW	chromosome mapping; tissue typing; preventive medicine; pharmacogenomic.	
OS	Homo sapiens.	
XX	WO200299116-A2.	
XX	12-DEC-2002.	
XX	04-JUN-2002; 2002WO-US017428.	
PR	04-JUN-2001; 2001US-0295607P.	
PR	04-JUN-2001; 2001US-0295661P.	
PR	06-JUN-2001; 2001US-0296404P.	
PR	06-JUN-2001; 2001US-0296418P.	
PR	14-JUN-2001; 2001US-0298285P.	
PR	15-JUN-2001; 2001US-0298556P.	
PR	21-JUN-2001; 2001US-0299949P.	
PR	26-JUN-2001; 2001US-0300883P.	
PR	28-JUN-2001; 2001US-0301550P.	
13-AUG-2001; 2001US-0311972P.		
27-AUG-2001; 2001US-0315071P.		
29-AUG-2001; 2001US-0315609P.		
14-SEP-2001; 2001US-0322933P.		
17-SEP-2001; 2001US-0322706P.		
14-DEC-2001; 2001US-0341186P.		
28-FEB-2002; 2002US-0361189P.		
12-MAR-2002; 2002US-0363673P.		
12-MAR-2002; 2002US-0363876P.		
03-JUN-2002; 2002US-00363676.		
(CURA-) CURAGEN CORP.		
PI	Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;	
PI	Gangoli RA, Gerlach VL, Gorman L, Guo X, Hjal T, Kekuda R, Li L;	
PI	Maddougall JR, Malyankar UM, Millet I, Padigaru M, Patturajan M;	
PI	Pena CE, Rastelli L, Shimkets RA, Stone DU, Spytek KA, Vernet CAM;	
PI	Voss EZ, Zerhusen BD;	
XX	WPI: 2003-140627/13.	
DR	N-PSDB; ACDD03623.	
XX	New NOVX polypeptides and nucleic acids, useful for preventing or	
PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,	
PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or	
PT	pharmacogenomics.	
XX	Claim 1; Page 99; 332pp; English.	
XX	The invention describes an isolated polypeptide (I) comprising any of 27	
CC	118-961 residue amino acid sequences, given in the specification, a	
CC	mature form of them, a sequence that is at least 95 % identical to them,	
CC	or a sequence having one or more conservative substitutions in them. The	
CC	polypeptide is useful in manufacturing a medicament for treating a	
CC	syndrome associated with a human disease selected from a pathology	
CC	associated with the polypeptide. The NOVX polypeptides, polynucleotides	
CC	and antibodies are useful in treating or preventing NOVX-associated	
CC	disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune	
CC	diseases, Alzheimer's disease, infections, multiple sclerosis, cancer-	
CC	associated cachexia, and other wasting disorders associated with chronic	
CC	diseases. The nucleic acids and polypeptides may also be used as targets	
CC	for the identification of small molecules that modulate or inhibit e.g.	
CC	neurogenesis, cell differentiation, cell proliferation, haematopoiesis,	
CC	wound healing and angiogenesis, in gene therapy, in generation of	
CC	antibodies that bind immunospecifically to NOVX substances for use in	
CC	therapeutic or diagnostic methods. The nucleic acids are further used as	
CC	hybridisation probes, in chromosome mapping, tissue typing, preventive	
CC	medicine, and pharmacogenomics. The polypeptides are also useful as	
CC	vaccines. This is the amino acid sequence of a novel human G-protein	
CC	coupled receptor related protein NOV	
XX	Sequence 582 AA;	
QY	373	MCVPLVAVVCTCLOQLEGGGLLQTSRTTAVVYLLSLMQPKGAPRLQPPNQ 432
DB	1	MCVPLVAVVCTCLOQLEGGGLLQTSRTTAVVYLLSLMQPKGAPRLQPPNQ 60
QY	433	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE 492
DB	61	GLCSLAADGLWNQKILFEEDQRLKHGLDGEDVSAFLNMNIFQDKINCERYYSFIHLSFQE 120
QY	493	FFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFLATSRFLGLNETHSHLEKSL 552
DB	121	FFAAMYIILDEGGAGPDQDVTLLTEYAFSERSFLATSRFLGLNETHSHLEKSL 180
QY	553	CKVSPHIKMDLLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHFQIVVS 612
DB	181	CKVSPHIKMDLLQWIOSKQSDGSTLQOQSLEFFSCLYEIQEEFFIQOALSHFQIVVS 240

QY 613 NTASKMEHVVSSFLKRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPRTVLL 672
 DB 241 NTASKMEHVVSSFLKRCRQAQVHLHYGATYSADGEDRARCAGATLLVQLRPRTVLL 300
 QY 673 DAYSEHLAALCTNPNIELSLYRNALGSRGVKLLCGQLRHFNCKLQNL 722
 DB 301 DAYSEHLAALCTNPNIELSLYRNALGSRGVKLLCGQLRHFNCKLQNL 350

RESULT 9

ABB77910
 ID ABB77910 standard; protein; 449 AA.

XX AC ABB77910;

XX DT 07-OCT-2002 (first entry)

XX Amino acid sequence of human leucine-rich repeat protein HLRBEM1.

XX Human; leucine-rich repeat; HLRBEM1; proliferative disorder;
 XX immune condition; apoptosis; signal transduction; autoimmune disease;
 XX haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
 XX cardiovascular disease; neurological disease; pheromone;
 XX pulmonary disease; chronic obstructive pulmonary disease;
 XX allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
 XX haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
 XX inflammatory disorder; systemic lupus erythematosus;
 XX cardiovascular disease; cancer.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers
 FT Domain 144..159
 FT /note= "transmembrane domain"

PN WO200252011-A2.

XX PD 04-JUL-2002.

XX PF 20-DEC-2001; 2001WO-US049740.

XX PR 22-DEC-2000; 2000US-0257773P.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Feder J, Ramanathan C, Mintier G;

XX DR WPI; 2002-566676/60.

XX DR N-PSDB; ABL59333.

XX New HLRBEM1 nucleic acids for preventing, treating or ameliorating e.g.
 FT proliferative disorders, immune conditions, a disorder related to
 FT aberrant apoptosis modulation or developmental disorders.

XX PS Claim 20; Fig 1A-E; 371pp; English.

XX The present sequence represents a human leucine-rich repeat containing
 CC protein, designated HLRBEM1. HLRBEM1 polypeptides and polynucleotides are
 CC useful for preventing, treating or ameliorating a medical condition such
 CC as a proliferative disorder, immune condition, or a disorder related to
 CC aberrant apoptosis modulation, either directly or indirectly, and in
 CC modulating signal transduction activity in various cells, tissue and
 CC organisms. They are also useful for treating, preventing, or diagnosing
 CC diseases of haematopoietic cells, autoimmune disease, graft-versus-host
 CC disease, allergic conditions (e.g. asthma), cardiovascular disorders, and
 CC neurological diseases, and for increasing the organisms' ability to
 CC synthesize and/or release pheromones. The polypeptide may also be used in
 CC treating, preventing or ameliorating pulmonary disease (e.g. chronic
 CC obstructive pulmonary disease, allergic rhinitis, or bronchial
 CC hyperresponsiveness), reproductive disease, haematopoietic disease,
 CC platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious
 CC disorders (e.g. innate immunity to bacterial pathogens, or adaptive

CC immune response), immune and inflammatory disorders (e.g. systemic lupus
 CC erythematosus), cardiovascular diseases and cancers. HLRBEM1 nucleic
 CC acids may further be used in chromosome identification or mapping, as a
 CC chromosome marker, as molecular weight markers, as diagnostic probes, in
 CC gene therapy, in raising anti-DNA antibodies, or as antigens for
 CC eliciting immune responses

XX SQ Sequence 449 AA;

Query Match 24.5%; Score 254; DB 5; Length 449;

Best Local Similarity 100.0%; Pred. No. 2.1e-243;

Matches 254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 228 MNQSAATECSMDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDELKPSFHDPOGPNCLCW 287

DB 1 MNQSAATECSMDLIFSCWPEPSAPLQELIRVPERLLFIIDGFDELKPSFHDPOGPNCLCW 60

QY 288 EEKRPTELLNSLRKKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSAEKREYF 347

DB 61 EEKRPTELLNSLRKKLLPELSLLITTRPTALEKHLRLEHPRHVEILGFSAEKREYF 120

QY 348 KYFHNAAEQAGQVFNVRDNEPLFTMCVFPLVCVVVCTCQQQLEGGLRQTSRTTAV 407

DB 121 KYFHNAAEQAGQVFNVRDNEPLFTMCVFPLVCVVVCTCQQQLEGGLRQTSRTTAV 180

QY 408 YMYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHLGDGVSAF 467

DB 181 YMYLLSLMQPKGAPRLQPPNQRGLCSLAADGLWNQKILFPEQDLRKHLGDGVSAF 240

QY 468 LNNNIFQKDINCER 481

DB 241 LNNNIFQKDINCER 254

RESULT 10

AAE21064

ID AAE21064 standard; protein; 287 AA.

XX AC AAE21064;

XX DT 01-JUL-2002 (first entry)

XX rno (upregulated by nitric oxide)-3 protein.

XX Upregulated by nitric oxide; NO; rno-3; differentiation; death; cancer;

XX acute myeloid leukaemia; AML; cytostatic.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT Misc-difference 249

FT /label= Unknown

FT /note= "This residue is represented in the specification

FT as O"

FT Misc-difference 256

FT /label= Unknown

FT /note= "This residue is represented in the specification

FT as O"

XX US2001029033-A1.

XX PD 11-OCT-2001.

XX PF 06-MAR-2001; 2001US-00799983.

XX PR 06-MAR-2000; 2000US-0186971P.

XX PA (SHAM/) SHAMI P J.

XX PA (PARK/) PARKER C J.

XX Shami PJ, Parker CJ;

XX WPI; 2002-009982/01.

XX Novel polypeptide regulated by nitric oxide useful for inducing acute
PT myeloid leukemia cell differentiation and apoptosis.
XX
XX
XX Claim 29; Page; 23pp; English.
XX
XX The present invention relates to novel genes which are upregulated by
CC nitric oxide (NO), designated as rno and their corresponding proteins.
CC The invention also relates to the isolation and characterisation of three
CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
CC are useful for inducing differentiation and death in a cancer cell.
CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
CC expression of a rho gene comprising the nucleotides coding for rno
CC protein. The present sequence is rno-3 protein. Note: This sequence is
CC not shown in the specification but is derived from rno-1 protein shown as
CC SEQ ID NO: 2 in figure 7B of the specification
XX
XX Sequence 287 AA;

Query Match 21.0%; Score 217; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 9.5e-207; Mismatches 0; Gaps 0;
Matches 217; Conservative 0; Indels 0; Gaps 0;
QY 723 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVFPFMMLLCEGLRHPQCRLOMIOLRK 782
Db 32 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVFPFMMLLCEGLRHPQCRLOMIOLRK 91
QY 783 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 842
Db 92 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 151
QY 843 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTQTLRLGICRLGSA 902
Db 152 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTQTLRLGICRLGSA 211
QY 903 ACEGLSVVLOANHNRLRELDLSFNDLGDWGLWLLAEG 939
Db 212 ACEGLSVVLOANHNRLRELDLSFNDLGDWGLWLLAEG 248

RESULT 11
AAE21062
ID AAE21062 standard; protein; 344 AA.
XX
AC AAE21062;
XX
DT 01-JUL-2002 (first entry)
XX
DE rno (upregulated by nitric oxide)-1 protein.
XX
XX Upregulated by nitric oxide; NO; rno-1; differentiation; death; cancer;
KW acute myeloid leukaemia; AML; cytostatic.
XX
OS Unidentified.
XX
XX Key Location/Qualifiers
PH Misc-difference 249
FT /label= Unknown
FT /note= "This residue is represented in the specification
FT as O"
FT Misc-difference 256
FT /label= Unknown
FT /note= "This residue is represented in the specification
FT as O"
XX
XX US2001029033-A1.
XX
XX 11-OCT-2001.
XX
XX 06-MAR-2001; 2001US-00799983.
XX
XX 06-MAR-2000; 2000US-0186971P.
XX

PA (SHAM/) SHAMI P J.
XX (PARK/) PARKER C J.
XX
PI Shami PJ, Parker CJ;
XX
XX MPI; 2002-009982/01.
XX
XX Novel polypeptide regulated by nitric oxide useful for inducing acute
PT myeloid leukemia cell differentiation and apoptosis.
XX
XX Claim 29; Fig 7B; 23pp; English.
XX
XX The present invention relates to novel genes which are upregulated by
CC nitric oxide (NO), designated as rno and their corresponding proteins.
CC The invention also relates to the isolation and characterisation of three
CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
CC are useful for inducing differentiation and death in a cancer cell.
CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
CC expression of a rho gene comprising the nucleotides coding for rno
CC protein. The present sequence is rno-1 protein
XX
XX Sequence 344 AA;

Query Match 21.0%; Score 217; DB 5; Length 344;
Best Local Similarity 100.0%; Pred. No. 1.1e-206;
Matches 217; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 723 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVFPFMMLLCEGLRHPQCRLOMIOLRK 782
Db 32 LKCRISSSACEDLSAALIANKNLTRMDLSGNGVFPFMMLLCEGLRHPQCRLOMIOLRK 91
QY 783 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 842
Db 92 COLESGACQEMASVLTGNPHLVELDTGNALEDGLRLCCGLRHPVCRILTLMKICRL 151
QY 843 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTQTLRLGICRLGSA 902
Db 152 TAAACDELASTLSVNSQSLRELDLSLNEGLDGLVLLCEGLRHPCKLTQTLRLGICRLGSA 211
QY 903 ACEGLSVVLOANHNRLRELDLSFNDLGDWGLWLLAEG 939
Db 212 ACEGLSVVLOANHNRLRELDLSFNDLGDWGLWLLAEG 248

RESULT 12
ADE36456
ID ADE36456 standard; protein; 190 AA.
XX
AC ADE36456;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human PAN6 NB-ARC domain seq id 63.
XX
XX cytostatic; immunosuppressive; vulnary; antinflammatory; vasotropic;
KW anti-allergic; antiulcer; dermatological; cerebroprotective; cardiac;
KW antiparkinsonian; neurotropic; neuroprotective; anti-HIV; gene therapy;
KW NPKappaB activation inhibitor; PAAD domain containing polypeptide;
KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
KW apoptosis; NPKappaB induction; cytokine processing;
KW cytokine receptor signaling caspase-mediated proteolysis;
KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
KW ulcerative colitis; graft versus host disease; stroke; heart failure;
KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
KW cancer therapy; PAAD domain family; human; PAN6; NB-ARC domain.
XX

OS Homo sapiens.
XX US2003077699-A1.
PN XX
PD XX
PP 24-APR-2003.
XX
PF 25-SEP-2001; 2001US-00965621.
XX
PR 26-SEP-2000; 2000US-00671760.
PR 26-SEP-2000; 2000US-0367367F.
XX
PA (REED/) REED J C.
PA (GODZ/) GODZIK A.
PA (CHUZ/) CHU Z.
PA (PAWL/) PAWLOWSKI K.
PA (FIOR/) FIORENTINO L.
PA (ARIZ/) ARIZA M E.
PA (STEH/) STEHLIK C.
XX
XX Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
PI Stehlik C;
XX
XX WPI; 2002-471256/50.
XX
XX Novel isolated PAD domain containing polypeptide useful for inducing
PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
PT therapy for treating cancer.
XX
XX Claim 22; SEQ ID NO 63; 93pp; English.
XX
XX The invention describes an isolated PAD domain containing polypeptide
CC (I) comprising 80% identity to the amino acid sequence of PAD and
CC nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck
CC -like protein containing a caspase recruitment domain (ASC)-2 fully
CC defined in specification, where (I) is biologically active. (I) is useful
CC for identifying a (I)-associated polypeptide, an agent altering that
CC association and agents that modulate PAD domain mediated inhibition of
CC nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful
CC for identifying an agent that modulates the activity of the NB-ARC domain
CC of (I). (I) or its functional fragments is useful in altering cellular or
CC biochemical processes such as apoptosis, NFkappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
CC -Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (I) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer
CC therapy. This is the amino acid sequence of a human PAD and nucleotide
CC binding protein PAN6 NB-ARC domain which demonstrates nucleotide binding
CC and facilitates the binding to another NB-ARC containing protein.
XX
SQ Sequence 190 AA;
Query Match 18.4%; Score 190; DB 5; Length 190;
Best Local Similarity 100.0%; Pred. No. 4.9e-180;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
183 PRTVWQGAAGIGKSLAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSMQDLIF 242
|||||
1 PRTVWQGAAGIGKSLAHKVMLDWADGKLFQGRFDYLFYINCREMNQATSCSMQDLIF 60
243 SCWPESAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLCWEEKPTELLNSLIR 302
|||||
61 SCWPESAPLOELIRYPERLLFIIDGFDLKPSPHDPQGPWCLCWEEKPTELLNSLIR 120
|||||
303 KKLLPELSLLITRPTALEKHLLEPHRVEILGFSAEERKEYFYKYPHNAEQGVFN 362
|||||

DB 121 KKLLPELSLLITRPTALEKHLLEPHRVEILGFSAEERKEYFYKYPHNAEQGVFN 180
QY 363 YVRDNEPLFT 372
|||||
DB 181 YVRDNEPLFT 190
RESULT 13
ADC32968
ID ADC32968 standard; protein; 197 AA.
XX
AC ADC32968;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3050.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnery;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 19.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
XX
XX N-PSDB; ADC32201.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Example 2; SEQ ID NO 3050; 1185pp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; a method of detecting
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are

CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a human contig-
 CC encoded polypeptide sequence used in an example of the invention. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 197 AA;

Query Match 18.4%; Score 190; DB 7; Length 197;
 Best Local Similarity 100.0%; Pred. No. 5.1e-180;
 Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 693 SLRYNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 752
 Db 1 SLRYNALGSRGVKLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 60
 QY 753 GNGVGFPGMMLLCEGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 812
 Db 61 GNGVGFPGMMLLCEGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 120
 QY 813 LEDLGLRLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 872
 Db 121 LEDLGLRLLCQGLRHPNCKLQNLKRCRISSACEDLSAALIANKNLTRMDLS 180
 QY 873 LGVLLLCCEGL 882
 Db 181 LGVLLLCCEGL 190

RESULT 14
 AAE21063
 ID AAE21063 standard; protein; 287 AA.
 AC AAE21063;
 XX
 XX 01-JUL-2002 (first entry)
 DE rno (upregulated by nitric oxide)-2 protein.
 XX
 XX Upregulated by nitric oxide; NO; rno-2; differentiation; death; cancer;
 KW acute myeloid leukaemia; AML; cytostatic.
 XX
 XX Unidentified.
 OS
 XX US2001029033-A1.
 PN
 XX 11-OCT-2001.
 PD
 XX 06-MAR-2001; 2001US-00799983.
 PF
 XX 06-MAR-2001; 2000US-0186971P.
 PR
 XX (SHAM/) SHAMI P J.
 PA (PARK/) PARKER C J.
 XX
 XX Shami PJ, Parker CJ;
 PI
 XX WPI; 2002-009982/01.
 DR
 XX Novel polypeptide regulated by nitric oxide useful for inducing acute
 PT myeloid leukemia cell differentiation and apoptosis.
 CC
 XX Claim 29; Page; 23pp; English.
 PS
 XX The present invention relates to novel genes which are upregulated by
 CC nitric oxide (NO), designated as rno and their corresponding proteins.
 CC The invention also relates to the isolation and characterisation of three

CC isoforms of rno gene, rno-1, rno-2, rno-3. rno proteins of the invention
 CC are useful for inducing differentiation and death in a cancer cell.
 CC Diagnosing acute myeloid leukaemia (AML) comprises detecting the
 CC expression of a rno gene comprising the nucleotides coding for rno
 CC protein. The present sequence is rno-2 protein. Note: This sequence is
 CC not shown in the specification but is derived from rno-1 protein shown as
 CC SEQ ID NO: 2 in figure 7B of the specification
 XX
 SQ Sequence 287 AA;

Query Match 16.6%; Score 172; DB 5; Length 287;
 Best Local Similarity 100.0%; Pred. No. 5.9e-162;
 Matches 172; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 723 LKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPNCKLQNLK 782
 Db 32 LKRCRISSACEDLSAALIANKNLTRMDLSGNGVGFPGMMLLCEGLRHPNCKLQNLK 91
 QY 783 COLESGACQEMASVLGTNPHLVLDLTGNALEDGLRLCQGLRHPNCKLQNLK 842
 Db 92 COLESGACQEMASVLGTNPHLVLDLTGNALEDGLRLCQGLRHPNCKLQNLK 151
 QY 843 TAAACDELASTLSVNQSLRELDLSLNEGLDGLVLLCEGLRHPNCKLQNLK 894
 Db 152 TAAACDELASTLSVNQSLRELDLSLNEGLDGLVLLCEGLRHPNCKLQNLK 203

RESULT 15
 ABG04570
 ID ABG04570 standard; protein; 168 AA.
 AC ABG04570;
 XX
 XX 13-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #4561.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 XX Homo sapiens.
 OS
 XX WO200175067-A2.
 PN
 XX 11-OCT-2001.
 PD
 XX 30-MAR-2001; 2001WO-US008631.
 PF
 XX 31-MAR-2000; 2000US-00540217.
 PR
 XX 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 XX Drmanac RT, Liu C, Tang YT;
 PI
 XX WPI; 2001-639362/73.
 DR
 XX N-PSDB; AAS68757.
 XX
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 XX Claim 20; SEQ ID NO 34929; 103pp; English.

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a

CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 168 AA;

Query Match 12.2%; Score 126; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 2.2e-116;
 Matches 126; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 539 LNEETRSHLEKSLCWKSPHIKMDLLOWIQSKAQSDGSTLOQGSLEFFSCLYEIOEEF 598
 Db 12 LNEETRSHLEKSLCWKSPHIKMDLLOWIQSKAQSDGSTLOQGSLEFFSCLYEIOEEF 71
 QY 599 IQQALSHFQVIVVSNIAKMEHNVSSFCCLKRCRSAQVHLHYGATYSADGEDRARCAGAH 658
 Db 72 IQQALSHFQVIVVSNIAKMEHNVSSFCCLKRCRSAQVHLHYGATYSADGEDRARCAGAH 131
 QY 659 TLLVQL 664
 Db 132 TLLVQL 137

Search completed: July 30, 2004, 14:02:48
 Job time : 60 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 14:00:43 ; Search time 21 Seconds
(without alignments)
4740.868 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 1035
Sequence: 1 MLRTAGRGGLCRSLTYLEEL.....MTHSRALALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR 78: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9	0.9	1192	T17255	hypothetical prote
2	8	0.8	87	AC0434	Bola-like protein
3	8	0.8	122	S60262	corticotensin - ra
4	8	0.8	123	D71922	probable flagellar
5	8	0.8	123	H64592	flagellar switch p
6	8	0.8	313	B95264	probable LysR-fam
7	8	0.8	360	S68209	sds22 protein homo
8	8	0.8	456	S20597	ribonuclease inhib
9	8	0.8	456	A31857	ribonuclease inhib
10	8	0.8	461	A31858	ribonuclease angio
11	8	0.8	475	1 Z4BPT4	helicase (EC 3.6.1
12	8	0.8	492	F86263	hypothetical prote
13	8	0.8	512	T38422	probable chromatin
14	8	0.8	611	A11941	hypothetical prote
15	8	0.8	661	I56258	RP105 - mouse
16	8	0.8	978	E96787	protein T4012.5 [i
17	8	0.8	1784	C96615	hypothetical prote
18	7	0.7	36	B44400	myosin heavy chain
19	7	0.7	50	B82279	hypothetical prote
20	7	0.7	54	S31533	ribulose-bisphosph
21	7	0.7	89	H82166	hypothetical prote
22	7	0.7	100	D41608	hima protein - rho
23	7	0.7	100	T49857	hypothetical prote
24	7	0.7	102	1 W4WL31	E4 protein - human
25	7	0.7	116	S62513	hypothetical 13.8K
26	7	0.7	121	A84083	hypothetical prote
27	7	0.7	135	F97942	hypothetical prote
28	7	0.7	141	T75252	hypothetical prote
29	7	0.7	142	A81635	hypothetical prote

30	7	0.7	154	2	C81579	conserved hypotet
31	7	0.7	159	2	H75391	hypothetical prote
32	7	0.7	171	2	D72090	hypothetical prote
33	7	0.7	171	2	F86532	hypothetical prote
34	7	0.7	179	2	F75519	Holliday junction
35	7	0.7	192	2	T06284	hypothetical prote
36	7	0.7	194	2	JC5260	progesterone membr
37	7	0.7	196	2	AI0555	4-methyl-5(b-hydro
38	7	0.7	197	2	H86741	conserved hypotet
39	7	0.7	198	1	H64771	hydroxymethylpyrim
40	7	0.7	198	2	F90688	hypothetical prote
41	7	0.7	198	2	B85539	hypothetical prote
42	7	0.7	208	2	AE2806	glutathione S-tran
43	7	0.7	208	2	D97585	probable DNA-3-met
44	7	0.7	215	2	G98285	DNA-3-methyladenin
45	7	0.7	218	2	AB2998	

ALIGNMENTS

RESULT 1

T17255
hypothetical protein DKFp58601822.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
R:Koehrer, K.; Beyer, A.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A:Reference number: Z18722
C:Accession: T17255
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1192 <KOE>
A:Cross-references: EMBL:AL117470
A:Experimental source: adult uterus; clone DKFp58601822
C:Genetics:
A:Note: DKFp58601822.1

Query Match 0.9%; Score 9; DB 2; Length 1192;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 QGAAGIGKS 197
DB 52 QGAAGIGKS 60

RESULT 2

AC0434
Bola-like protein [imported] - Yersinia pestis (strain CO92)
C:Species: Yersinia pestis
C>Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
C:Accession: AC0434
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-87 <KUR>
C:Genetics:
A:Gene: YPO3570
C:Superfamily: bola protein

Query Match 0.8%; Score 8; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 SHFQIVV 611
 Db 29 SHFQIVV 36

RESULT 3

S60262
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 10-Apr-1996 #sequence_revision 19-Apr-1996 #text_change 20-Jan-2003
 C:Accession: S60262
 R:Vaughan, J.; Donaldson, C.; Bittencourt, J.; Perrin, M.H.; Lewis, K.; Sutton, S.; Chan
 Nature 378, 287-292, 1995
 A:Title: Urocortin, a mammalian neuropeptide related to fish urotensin I and to corticot
 A:Reference number: S60262; MUID:96069764; PMID:7477349
 A:Accession: S60262
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-122 <VAN>
 A:Cross-references: EMBL:U33935; NID:G1065908; PIDN:AAA87566.1; PID:G1065909
 C:Superfamily: corticotropin-releasing factor receptor type 1; diuretic hormone homology
 F;80-120/Domain: diuretic hormone homology <DHH>

Query Match 0.8%; Score 8; DB 2; Length 122;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 660 LLVQLRPE 667
 Db 15 LLVQLRPE 22

RESULT 4

D71922
 Probable flagellar motor switch protein - Helicobacter pylori (strain J99)
 C:Species: Helicobacter pylori
 A:Variety: strain J99
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 26-Aug-1999
 R:Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.L.; Brown, E.D.; Doig, P.C.; Smith, D.R.;
 Ives, C.; Gibson, R.; Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Vovis, G.F.;
 Nature 397, 176-180, 1999
 A:Title: Genomic sequence comparison of two unrelated isolates of the human gastric path
 A:Reference number: A71800; MUID:99120557; PMID:9923682
 A:Accession: D71922
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-123 <ARN>
 A:Cross-references: GB:AE001485; GB:AE001439; NID:G4155057; PIDN:AAD06097.1; PID:G415506
 A:Experimental source: strain J99
 C:Genetics:
 A:Gene: flin
 C:Superfamily: flagellar motor switch protein

Query Match 0.8%; Score 8; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSTYLEEL 20
 Db 27 LSTYLEEL 34

RESULT 5

H64592
 Flagellar switch protein - Helicobacter pylori (strain 26695)
 C:Species: Helicobacter pylori
 C:Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 26-Aug-1999
 C:Accession: H64592
 R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.;
 Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khairak, H.G.; Glodek, A.; McKen
 son, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L.
 Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.
 A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
 A:Reference number: A64520; MUID:97394467; PMID:9252185
 A:Accession: H64592
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-123 <TOM>
 A:Cross-references: GB:AE000571; GB:AE000511; NID:G2313686; PIDN:AAD07636.1; PID:G2313686
 C:Superfamily: flagellar motor switch protein

Query Match 0.8%; Score 8; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 5.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 LSTYLEEL 20
 Db 27 LSTYLEEL 34

RESULT 6

B95264
 Probable lysozyme family protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasma
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 27-Nov-2001
 C:Accession: B95264
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
 Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: B95264
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-313 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK64676.1; PID:G14523075; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmaid pSYMA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
 Pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
 heubalt, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMA0039
 A:Genome: plasmid
 C:Superfamily: conserved hypothetical protein H11364

Query Match 0.8%; Score 8; DB 2; Length 313;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 919 ELDLSFND 926
 Db 136 ELDLSFND 143

RESULT 7

S68209
 Sds22 protein homolog - human
 C:Species: Homo sapiens (man)
 C:Date: 15-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
 C:Accession: S68209
 R:Renouf, S.; Beullens, M.; Wera, S.; Van Eynde, A.; Sikela, J.; Stalmans, W.; Bollen,
 FEBS Lett. 375, 75-78, 1995
 A:Title: Molecular cloning of a human polypeptide related to yeast sds22, a regulator
 A:Reference number: S68209; MUID:96087087; PMID:7498485
 A:Accession: S68209
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-360 <REN>

A;Cross-references: EMBL:Z50749; NID:gl085027; PIDN:CAA90626.1; PID:gl085028
 C;Superfamily: decorin; leucine-rich alpha-2-glycoprotein repeat homology; proteoglycan

Query Match 0.8%; Score 8; DB 2; Length 360;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 858 QSLRELDL 865
 DB 120 QSLRELDL 127
 |||||

RESULT 8
 S20597
 Ribonuclease inhibitor - rat
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 21-Jan-2000
 C;Accession: S20597
 R;Kawanomoto, M.; Motojima, K.; Sasaki, M.; Hattori, H.; Goto, S.
 Biochim. Biophys. Acta 1129, 335-338, 1992
 A;Title: cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue distribution
 A;Reference number: S20597; MUID:92162755; PMID:1536887
 A;Accession: S20597
 A;Molecule type: mRNA
 A;Residues: 1-456 <Xaw>
 A;Cross-references: EMBL:X62528; NID:g57670; PIDN:CAA44398.1; PID:g57671
 C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 0.8%; Score 8; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 CRLRLTDL 837
 DB 251 CRLRLTDL 258
 |||||

RESULT 9
 A31857
 Ribonuclease inhibitor, hepatic - pig
 C;Species: Sus scrofa domestica (domestic pig)
 C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 21-Jan-2000
 C;Accession: A31857; A35830
 R;Hofsteenge, J.; Kieffer, B.; Matthies, R.; Hemmings, B.A.; Stone, S.R.
 Biochemistry 27, 8537-8544, 1988
 A;Title: Amino acid sequence of the ribonuclease inhibitor from porcine liver reveals th
 A;Reference number: A31857; MUID:89118268; PMID:3219361
 A;Accession: A31857
 A;Molecule type: protein
 A;Residues: 1-456 <HOF>
 R;Vicentini, A.M.; Kieffer, B.; Matthies, R.; Meyhack, B.; Hemmings, B.A.; Stone, S.R.;
 Biochemistry 29, 8827-8834, 1990
 A;Title: Protein chemical and kinetic characterization of recombinant porcine ribonuclea
 A;Reference number: A35830; MUID:91104783; PMID:2271559
 A;Accession: A35830
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 82-456 <VIC>
 A;Cross-references: GB:M58700; NID:G164638; PIDN:AAA63448.1; PID:G164639; GB:J02925
 C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology
 C;Keywords: liver
 F;280-304/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR>

Query Match 0.8%; Score 8; DB 2; Length 456;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SURELDLS 866
 DB 394 SURELDLS 401
 |||||

RESULT 10

A31858

Ribonuclease-angiogenesis inhibitor - human

N;Alternate names: ribonuclease inhibitor, placental

C;Species: Homo sapiens (man)

C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 19-May-2000

C;Accession: A31858; S02012; S23933; S48636; T47188

R;Lee, F.S.; Fox, E.A.; Zhou, H.M.; Strydom, D.J.; Vallee, B.L.

Biochemistry 27, 8545-8553, 1988

A;Title: Primary structure of human placental ribonuclease inhibitor.

A;Reference number: A31858; MUID:89118269; PMID:3219362

A;Accession: A31858

A;Molecule type: mRNA

A;Residues: 1-461 <LEE>

A;Cross-references: GB:M22414; NID:G186360; PIDN:AAAS9130.1; PID:G307040

R;Schneider, R.; Schneider-Scherzer, E.; Thurnher, M.; Auer, B.; Schweiger, M.

EMBO J. 7, 4151-4156, 1988

A;Title: The primary structure of human ribonuclease/angiogenesis inhibitor (RAI) disclose

A;Reference number: S02012; MUID:89210799; PMID:3243277

A;Accession: S02012

A;Molecule type: mRNA

A;Residues: 1-422, 'SE', 425-461 <SCH>

A;Cross-references: EMBL:X13973; NID:G35843; PIDN:CAA32151.1; PID:G35844

A;Note: part of this sequence, including the carboxyl end of the mature protein, was cor

R;Crevel-Thieffry, I.; Cottrell, S.; Schuller, E.

Biochim. Biophys. Acta 1122, 107-112, 1992

A;Title: Characterisation of a tryptic peptide from human placental ribonuclease inhibit

A;Reference number: S23933; MUID:92338217; PMID:1633192

A;Accession: S23933

A;Molecule type: protein

A;Residues: 174-195 <CRE>

R;Nadano, D.; Yasuda, T.; Takeshita, H.; Uchida, K.; Kishi, K.

Arch. Biochem. Biophys. 312, 421-428, 1994

A;Title: Purification and characterization of human brain ribonuclease inhibitor.

A;Reference number: S48636; MUID:94311593; PMID:8037455

A;Accession: S48636

A;Status: preliminary

A;Molecule type: protein

A;Residues: 2-14 <NAD>

R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24374

A;Accession: T47188

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-461 <AAA>

A;Cross-references: EMBL:AL161967

A;Experimental source: adult testis; clone DKFZp434K249

C;Genetics:

A;Gene: GDB:RNH

A;Cross-references: GDB:125274; OMIM:173320

A;Map position: lip15.5-lip15.5

A;Note: DKFZp434K249.1

C;Superfamily: ribonuclease inhibitor; leucine-rich alpha-2-glycoprotein repeat homology

C;Keywords: blocked amino end; duplication

F;34-433/Region: leucine-rich 57-residue repeats

F;57-81/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR1>

F;114-138/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR2>

F;174-195/Region: inhibitory

F;228-252/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR3>

F;285-309/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR4>

F;399-423/Domain: leucine-rich alpha-2-glycoprotein repeat homology <LRR5>

Query Match 0.8%; Score 8; DB 2; Length 461;

Best Local Similarity 100.0%; Pred. No. 18;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SURELDLS 866

DB 399 SURELDLS 406

|||||

RESULT 11

ZABPT4
 helicase (EC 3.6.1.-) primase chain - phase T4
 N;Alternate names: gp 41
 C;Species: phage T4
 A;Note: host Escherichia coli
 C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 19-Jan-2001
 C;Accession: A04308
 R;Nakanishi, M.; Alberts, B.
 unpublished results 1985, cited by GenBank
 A;Reference number: A94456
 A;Accession: A04308
 A;Molecule type: DNA
 A;Residues: 1-475 <DNA>
 A;Cross-references: GB:K03113; NID:g215964; PIDN:AAA32553.1; PID:g215966
 R;Young, M.C.; Schultz, D.E.; Ring, D.; von Hippel, P.H.
 J. Mol. Biol. 235, 1447-1458, 1994
 A;Title: Kinetic parameters of the translocation of bacteriophage T4 gene 41 protein hel
 A;Reference number: A58645; MUID:94149695; PMID:8107085
 A;Contents: annotation; enzyme activity
 C;Comment: This protein has GTPase, GTPase and dATPase activities, and it is a helicase
 C;Genetics:
 A;Gene: 41
 A;Map position: 20.619-22.044
 A;Start codon: GUG
 C;Superfamily: phage T4 gene 41 protein
 C;Keywords: DNA binding; DNA repair; DNA replication; hydrolase; nucleotide binding; P-1
 F;197-204/Region: nucleotide-binding motif A (P-loop)

Query Match 0.8%; Score 8; DB 1; Length 475;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 433 GLCSLAAD 440
 |||||
 Db 206 GLCSLAAD 213

RESULT 12
 F86263
 hypothetical protein Flk23.23 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
 C;Accession: F86263
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712
 A;Accession: F86263
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-492 <STO>
 A;Cross-references: GB:AB051172; NID:g8698744; PIDN:AAF78502.1; GSPDB:GN00141
 C;Genetics:
 A;Map position: 1

Query Match 0.8%; Score 8; DB 2; Length 492;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 916 NIRELDS 923
 |||||
 Db 377 NIRELDS 384

RESULT 13
 T38422

probable chromatin assembly factor subunit - fission yeast (Schizosaccharomyces pombe)
 C;Species: Schizosaccharomyces pombe
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000
 C;Accession: T38422
 R;Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, September 1997
 A;Reference number: Z21748
 A;Accession: T38422
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-512 <OLI>
 A;Cross-references: EMBL:Z99126; PIDN:CAB16189.1; GSPDB:GN00066; SPDB:SPAC26H5.03
 A;Experimental source: strain 972h-; cosmid c26H5
 C;Genetics:
 A;Gene: SPDB:SPAC26H5.03
 A;Map position: 1
 A;Introns: 44/3; 79/1; 160/1
 C;Superfamily: unassigned WD repeat proteins; WD repeat homology

Query Match 0.8%; Score 8; DB 2; Length 512;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 589 CLYEIQEE 596
 |||||
 Db 195 CLYEIQEE 202

RESULT 14
 A11941
 hypothetical protein all1084 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
 C;Accession: A11941
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata,
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: A11941
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-611 <KUR>
 A;Cross-references: GB:BA000019; PIDN:BAB73041.1; PID:g17130430; GSPDB:GN00179
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all1084

Query Match 0.8%; Score 8; DB 2; Length 611;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 920 LDLSFNLD 927
 |||||
 Db 593 LDLSFNLD 600

RESULT 15
 RPI05 - mouse
 C;Species: Mus musculus (house mouse)
 C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 05-Nov-1999
 C;Accession: I56258
 R;Miyake, K.; Yamashita, Y.; Ogata, M.; Sudo, T.; Kimoto, M.
 J. Immunol. 154, 3333-3340, 1995
 A;Title: RPI05, a novel B cell surface molecule implicated in B cell activation, is a
 A;Reference number: I56258; MUID:95204928; PMID:7897216
 A;Accession: I56258
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: mRNA
 A;Residues: 1-661 <RES>
 A;Cross-references: GB:D37797; NID:g761711; PIDN:BAA07043.1; PID:g761712

C;Superfamily: leucine-rich alpha-2-glycoprotein repeat homology

Query Match 0.8%; Score 8; DB 2; Length 661;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 NLRELDLS 923
 |||||
 Db 371 NLRELDLS 378

Search completed: July 30, 2004, 14:04:47
 Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 14:00:08 ; Search time 18 Seconds
(without alignments)

2994.034 Million cell updates/sec

Title: US-10-781-294-24

Perfect score: 1035

Sequence: 1 MLRTAGDGLCRSLTYLEEL.....MTHSRLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	818	79.0	1061	1 NAL12 HUMAN	P59046 homo sapien
2	13	1.3	1033	1 C1S1 MOUSE	Q8r4b8 mus musculus
3	13	1.3	1034	1 C1S1 HUMAN	Q96p20 homo sapien
4	12	1.2	1200	1 NAL5 HUMAN	P59047 homo sapien
5	9	0.9	980	1 NAL7 HUMAN	Q8wx94 homo sapien
6	9	0.9	1062	1 NAL2 HUMAN	Q9nx02 homo sapien
7	9	0.9	1093	1 NAL4 HUMAN	Q86w24 homo sapien
8	9	0.9	1473	1 NAL1 HUMAN	Q9c000 homo sapien
9	8	0.8	122	1 UCNI RAT	P55090 rattus norv
10	8	0.8	456	1 RINI FIG	P10775 sus scrofa
11	8	0.8	456	1 RINI RAT	P29315 rattus norv
12	8	0.8	460	1 RINI HUMAN	P13489 homo sapien
13	8	0.8	460	1 RINI PANTR	Q8h2p9 pan troglod
14	8	0.8	475	1 VG41 BPT4	P04530 bacterioph
15	8	0.8	661	1 C180 MOUSE	Q62192 mus musculus
16	8	0.8	707	1 IF2 THETN	Q8ra37 thermonaer
17	8	0.8	843	1 NAL6 MOUSE	Q91w62 mus musculus
18	8	0.8	980	1 CLV1 ARATH	Q9SYG8 arabidopsis
19	8	0.8	4036	1 RRLP DUGBV	Q66431 dugbe virus
20	7	0.7	54	1 RBL MAGLI	P31193 magnolia li
21	7	0.7	99	1 IHFA RHCCA	P30787 rhodobacter
22	7	0.7	102	1 VE4 HPV11	P17384 human papil
23	7	0.7	116	1 YAP7 SCHPO	Q09861 schizosacch
24	7	0.7	179	1 RUVC DEIRA	Q9R775 delinococcus
25	7	0.7	193	1 PGCI FIG	Q95250 sus scrofa
26	7	0.7	194	1 PGCI MOUSE	O55022 mus musculus
27	7	0.7	194	1 PGCI RAT	P70580 rattus norv
28	7	0.7	195	1 ASC HUMAN	Q9ulz3 homo sapien
29	7	0.7	196	1 THIU ECOLI	Q46948 escherichia
30	7	0.7	196	1 THIU SALTY	P55880 salmonella
31	7	0.7	199	1 COX2 RHEAM	O03893 rhea americ
32	7	0.7	224	1 KCY THETN	Q8ra78 thermonaer
33	7	0.7	238	1 ZNUC_BUCAI	P57403 buchnera ap

RESULT 1

ID	NAL12_HUMAN	STANDARD;	PRT;	1061 AA.
AC	P59046;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	10-OCT-2003 (Rel. 42, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	NACHT-, LRR- and PYD-containing protein 12 (PYRIN-containing APAF1-like protein 7) (Monarch-1).			
GN	NALP12 OR PYPAF7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=22162427; PubMed=12019269;			
RA	Mang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,			
RA	Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;			
RT	"PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates			
RT	activation of NF-kappa B and caspase-1-dependent cytokine			
RT	processing.";			
RL	J. Biol. Chem. 277:29874-29880(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=22451042; PubMed=12563287;			
RA	Tschopp J., Martinon F., Burns K.;			
RT	"NALPs: a novel protein family involved in inflammation.";			
RL	Nat. Rev. Mol. Cell Biol. 4:95-104(2003).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).			
RC	TISSUE=Lymphoma;			
RA	Williams K.L., Linhoff M.W., Harton J.A., Ting J.P.Y.;			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Leukocyte;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,			
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,			
RA	Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;			
RT	"Generation and initial analysis of more than 15,000 full-length			

P11402 spinacia ol
Q9by71 homo sapien
P59034 mus musculus
P59035 rattus norv
P25637 saccharomyc
Q9veb5 drosophila
Q02760 rhodobacter
Q55752 synechocyst
P33574 petromyzon
P39261 bacterioph
Q9ttt8 trichosurus
Q9m518 arabidopsis

ALIGNMENTS

RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- FUNCTION: May mediate activation of CASP1 via ASC and promote
 CC activation of NF-kappa-B via IKK.
 CC -!- SUBUNIT: Binds to ASC with its DAPIN domain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=1; Synonyms=1;
 CC IsoId=P59046-1; Sequence=Displayed;
 CC Name=2; Synonyms=11;
 CC IsoId=P59046-2; Sequence=VSP_005524;
 CC Name=3; Synonyms=11;
 CC IsoId=P59046-3; Sequence=VSP_005523;
 CC -!- TISSUE SPECIFICITY: Detected only in peripheral blood leukocytes,
 CC predominantly in eosinophils and granulocytes, and at lower levels
 CC in monocytes.
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 8 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AY095146; AAM18227.1; -
 CC EMBL; AY154467; AAO18163.1; -
 CC EMBL; AY116204; AAM75142.1; -
 CC EMBL; AY116205; AAM75143.1; -
 CC EMBL; AY116206; AAM75144.1; -
 CC EMBL; BC028069; AAM28069.1; -
 CC Genew; HGNC:22938; NALP12.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR007091; LRR_RNinh.
 CC InterPro; IPR003590; LRR_RNinh_sub.
 CC InterPro; IPR007111; NACHT_NTPase.
 CC InterPro; IPR004020; PAAD_DAPIN_dom.
 CC Pfam; PF00560; LRR_2.
 CC Pfam; PF02758; PAAD_DAPIN; 1.
 CC SMART; SM00388; LRR_R1; 11.
 CC PROSITE; PS0824; DAPIN; 1.
 CC PROSITE; PS0837; NACHT; 1.
 CC KW ATP-binding; Leucine-rich repeat; Repeat; Alternative splicing.
 FT DOMAIN 1 95
 FT DAPIN.
 FT DAPIN 211 528
 FT REPEAT 712 735
 FT REPEAT 741 764
 FT REPEAT 769 791
 FT REPEAT 798 821
 FT REPEAT 826 849
 FT REPEAT 883 906
 FT REPEAT 940 968
 FT REPEAT 997 1020
 FT NP_BIND 217 224
 FT VARSPPLIC 862 973
 FT VARSPPLIC 976 1031
 FT CONFLICT 691 691 L -> LR (IN REF. 3).
 FT SEQUENCE 1061 AA; 120172 MW; 8C10AFA907C131B CRC64;
 Query Match 79.0%; Score 818; DB 1; Length 1061;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 938; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 97 DPQETRYDYVRKFRIMEDRNARLGCYNLSHRYTLLAVKHSNPMQVQQQLDITGRGH 156
 DB 124 DPQETRYDYVRKFRIMEDRNARLGCYNLSHRYTLLAVKHSNPMQVQQQLDITGRGH 183

QY 157 ARTVGHQASPIKIETLFEPEDEERPEPRPTVMQGAAGIGKSMIAHVMLDWADGKLPQGR 216
 DB 184 ARTVGHQASPIKIETLFEPEDEERPEPRPTVMQGAAGIGKSMIAHVMLDWADGKLPQGR 243
 QY 217 FDYLFYINCREMOSATECSMODLIFSCWPEPAPLOELIRVPERLLFIIDGDELKPSF 276
 DB 244 FDYLFYINCREMOSATECSMODLIFSCWPEPAPLOELIRVPERLLFIIDGDELKPSF 303
 QY 277 HDPOGPWCLCWEKRPTELLNSLIRKLLPELSLLITTPRTALEKHLRLEHPRHVEIL 336
 DB 304 HDPOGPWCLCWEKRPTELLNSLIRKLLPELSLLITTPRTALEKHLRLEHPRHVEIL 363
 QY 337 GFSEAEKKEYFYKYFHNAEQAGQVFNVRDNEPLFTMCVFVPLVCVVVCTCLOQLEGGL 396
 DB 364 GFSEAEKKEYFYKYFHNAEQAGQVFNVRDNEPLFTMCVFVPLVCVVVCTCLOQLEGGL 423
 QY 397 LROTSRTTAVVYMLYLLSMQKPGAPRLOPPNORGLCSLAADGLWNQKILFEEQDLRK 456
 DB 424 LROTSRTTAVVYMLYLLSMQKPGAPRLOPPNORGLCSLAADGLWNQKILFEEQDLRK 483
 QY 457 HGLDGEDVSFAFLNNNIFQKINCERYYSFIHLSPQEFFAAMYIILDEGEGAGDQDVTR 516
 DB 484 HGLDGEDVSFAFLNNNIFQKINCERYYSFIHLSPQEFFAAMYIILDEGEGAGDQDVTR 543
 QY 517 LLTEYAFSERSFLALTSRFLPGLLNBEHTRSHLEKSLCWKVSPIHMDLLOIQAQSDG 576
 DB 544 LLTEYAFSERSFLALTSRFLPGLLNBEHTRSHLEKSLCWKVSPIHMDLLOIQAQSDG 603
 QY 577 STLOOGLSEFPCLYEIQEEFIQOALSHFQVIVVSNIAKMEHMSVSPCLKRCRSQVL 636
 DB 604 STLOOGLSEFPCLYEIQEEFIQOALSHFQVIVVSNIAKMEHMSVSPCLKRCRSQVL 663
 QY 637 HLYGATYSADGEDRARCAGAHLLVQLRPTVLLDAYSEHAAALCTNPNIELSILYR 696
 DB 664 HLYGATYSADGEDRARCAGAHLLVQLRPTVLLDAYSEHAAALCTNPNIELSILYR 722
 QY 697 NALSGRGVLLCCQGLRHPNCKLNLKRCRISSEACEDLSAALIANKNLTRMDLSNGV 756
 DB 723 NALSGRGVLLCCQGLRHPNCKLNLKRCRISSEACEDLSAALIANKNLTRMDLSNGV 782
 QY 757 GFPGMMLLCEGLRHPQCLQMLQIRKQLESAGCQEMASVLGTNPHLVELDLTGNAEDL 816
 DB 783 GFPGMMLLCEGLRHPQCLQMLQIRKQLESAGCQEMASVLGTNPHLVELDLTGNAEDL 842
 QY 817 GLRLCCQGLRHPVCLRLTLKICRLTAAACDELASTLSVNQSLRELDLSNELGDLGVL 876
 DB 843 GLRLCCQGLRHPVCLRLTLKICRLTAAACDELASTLSVNQSLRELDLSNELGDLGVL 902
 QY 877 LLCEGLRHPCKLQTLRIGICRLGSAACEGLSVVLOAHNRLRELDLSNDLGDWGLWLLA 936
 DB 903 LLCEGLRHPCKLQTLRIGICRLGSAACEGLSVVLOAHNRLRELDLSNDLGDWGLWLLA 962
 QY 937 EGLQHPACRLQKMLDSCGLTAKACENLYFTLGINQTLTDLVLTNNALGDTGVRLCKRL 996
 DB 963 EGLQHPACRLQKMLDSCGLTAKACENLYFTLGINQTLTDLVLTNNALGDTGVRLCKRL 1022
 QY 997 SHPGCKLRLVLFGLMDLNKMTSHSLAALRVTKPYLDIGC 1035
 DB 1023 SHPGCKLRLVLFGLMDLNKMTSHSLAALRVTKPYLDIGC 1061
 RESULT 2
 CIS1 MOUSE
 ID CIS1 MOUSE STANDARD; PRT: 1033 AA.
 AC Q8RA48;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Cold autoinflammatory syndrome 1 protein homolog (Pyrin-containing
 DE APAF1-like protein 1) (Nast cell maturation inducible protein 1).
 GN CIA1 OR NALP3 OR PYPAF1 OR MMIG1.
 OS Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALE/cj;
 RA Kikuchi-Yoshita R., Koga K., Taketomi Y., Sugiki T., Saico T.,
 RA Ishii S., Hisada M., Suzuki-Nishimura T., Uchida M.K., Moon T.-C.,
 RA Chang H.-W., Sawada M., Inagaki N., Nagai H., Murakami M., Kudo I.;
 RT "Identification of inducible genes during in vitro maturation of mouse
 RT bone marrow-derived mast cells to connective tissue-type mast cells";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: May function as a potential inducer of apoptosis.
 CC Interacts selectively with apoptosis-associated specklike protein
 CC containing a CARD domain (ASC). This complex may function as an
 CC upstream activator of NF-kappaB signaling (By similarity).
 CC -!- SIMILARITY: Contains 1 DAPIN domain.
 CC -!- SIMILARITY: Contains 1 NACHT domain.
 CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF486632; AAL90874.1; -;
 DR MGD; MGI:2653833; C1asl.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR003590; LRR_RNinh_sub.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF00560; LRR; 2.
 DR Pfam; PF05729; NACHT; 1.
 DR Pfam; PF02758; PAAD_DAPIN; 1.
 DR SMART; SM00368; LRR_RL; 1.
 DR PROSITE; PS50824; DAPIN; 1.
 DR PROSITE; PS50837; NACHT; 1.
 DR Apoptosis; Repeat; Leucine-rich repeat.
 FT DOMAIN 1 91
 FT DOMAIN 216 532
 FT REPEAT 737 760
 FT REPEAT 794 817
 FT REPEAT 851 874
 FT REPEAT 880 903
 FT REPEAT 908 931
 FT REPEAT 937 964
 FT REPEAT 965 988
 FT REPEAT 1033 1033 AA; 118274 MW; 5924690966B12117 CRC64;
 SQ
 Query Match 1.3%; Score 13; DB 1; Length 1033;
 Best Local Similarity 100.0%; Pred. No. 0.00018;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 329 HPRHVEILGSEA 341 PRT; 1034 AA.
 Db 360 HPRHVEILGSEA 372
 RESULT 3
 C1SL_HUMAN STANDARD; PRT; 1034 AA.
 ID C1SL_HUMAN STANDARD; PRT; 1034 AA.
 AC Q96P20; O75434; Q8TCW0; Q8TEU9; Q8WXH9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cold autoimmune syndrome 1 protein (Cryopyrin) (NACHT-, LRR- and
 DE PYD-containing protein 3) (PRIN-containing APAF1-like protein 1)
 DE (Angiotensin/vasopressin receptor AII/AVP-like).
 GN C1AS1 OR NALP3 OR PYPAF1.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2); VARIANTS FCAS MET-198; VAL-439
 RP AND GLY-627, AND VARIANT MWS VAL-352.
 RX MEDLINE=21547523; PubMed=11687797;
 RA Hoffman H.M., Mueller J.L., Broide D.H., Wanderer A.A.,
 RA Kolodner R.D.;
 RT "Mutation of a new gene encoding a putative pyrin-like protein causes
 RT familial cold autoinflammatory syndrome and Muckle-Wells syndrome";
 RL Nat. Genet. 29:301-305(2001).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=21909508; PubMed=11786556;
 RA Manji G.A., Wang L., Geddes B.J., Brown M., Merriam S., Al-Garawi A.,
 RA Mak S., Lora J.M., Briskin M., Jurman M., Cao J., Distefano P.S.,
 RA Bertin J.;
 RT "PYPAF1: a PYRIN-containing APAF1-like protein that assembles with ASC
 RT and activates NF-kB";
 RL J. Biol. Chem. 277:11570-11575(2002).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3); VARIANT MWS MET-198, AND
 RP VARIANTS FCAS TRP-260 AND PRO-305.
 RX MEDLINE=22241234; PubMed=12355493;
 RA Aganna E., Martinon F., Hawkins P.N., Ross J.B., Swan D.C.,
 RA Booth D.R., Lachmann H.J., Gaudet R., Woo P., Feighery C.,
 RA Cotter F.E., Thome M., Hitman G.A., Teichmann J., McDermott M.F.;
 RT "Association of mutations in the NALP3/C1AS1/PYPAF1 gene with a broad
 RT phenotype including recurrent fever, cold sensitivity, sensorineural
 RT deafness, and AA amyloidosis";
 RL Arthritis Rheum. 45:2445-2452(2002).
 RN [4]
 RP SEQUENCE OF 391-1034 FROM N.A. (ISOFORM 1).
 RC TISSUE=Blood;
 RX MEDLINE=20499367; PubMed=11042152;
 RA Zhang Q.-H., Ye M., Wu X.Y., Ren S.X., Zhao M., Zhao C.J., Fu G.,
 RA Shen Y., Fan H.Y., Lu G., Zhong M., Xu X.R., Han Z.-G., Zhang J.W.,
 RA Tao J., Huang Q.-H., Zhou J., Hu G.X., Gu X.J., Chen S.-J., Chen Z.;
 RT "Cloning and functional analysis of cDNAs with open reading frames for
 RT 300 previously undefined genes expressed in CD34+ hematopoietic
 RT stem/progenitor cells";
 RL Genome Res. 10:1546-1560(2000).
 RN [5]
 RP VARIANTS FCAS MET-198, VARIANTS MWS ASN-303; MET-348; THR-439 AND
 RP ARG-569, AND VARIANT FCAS/MWS TRP-260.
 RX MEDLINE=21987640; PubMed=11992256;
 RA Dode C., Le Du N., Cuisset L., Letourneur F., Berthelot J.-M.,
 RA Vaudour G., Meyrier A., Watts R.A., Scott D.G.I., Nicholls A.,
 RA Granel B., Frances C., Garcier F., Edery P., Boulanguet S.,
 RA Domergues J.-P., Delpech M., Grateau G.;
 RT "New mutations of C1AS1 that are responsible for Muckle-Wells syndrome
 RT and familial cold urticaria: a novel mutation underlies both
 RT syndromes";
 RL Am. J. Hum. Genet. 70:1498-1506(2002).
 RN [6]
 RP VARIANTS CINCA ASN-303; SER-309; ARG-358; ASN-436; SER-573 AND
 RP THR-662, AND TISSUE SPECIFICITY.
 RX MEDLINE=22062556; PubMed=12032915;
 RA Feldmann J., Prieur A.-M., Quartier P., Berquin P., Certain S.,
 RA Cortis E., Teillac-Hamel D., Fischer A., de Saint Basile G.;
 RT "Chronic infantile neurological cutaneous and articular syndrome is
 RT caused by mutations in C1AS1, a gene highly expressed in
 RT polymorphonuclear cells and chondrocytes";
 RL Am. J. Hum. Genet. 71:198-203(2002).
 CC -!- FUNCTION: May function as a potential inducer of apoptosis.
 CC Interacts selectively with apoptosis-associated specklike protein
 CC containing a CARD domain (ASC). This complex may function as an
 CC upstream activator of NF-kappaB signaling.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=2;
 CC IsoId=Q96P20-1; Sequence=Displayed;


```

RT for early embryonic development.";
RL Hum. Reprod. 17:903-911(2002).
CC -!- TISSUE SPECIFICITY: Oocyte specific.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 13 leucine-rich (LRR) repeats.
CC -!- CAUTION: It is not obvious that this is the ortholog of mouse
CC Mater.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AY054986; AAL15549.1; -.
CC Genew; HGNC:21269; NALP5.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_RNinh.
CC InterPro; IPR003590; LRR_RNinh_sub.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF00560; LRR: 3.
CC Pfam; PF05729; NACHT; 1.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC SMART; SM00368; LRR_R1; 11.
CC PROSITE; PS50824; DAPIN; 1.
CC PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 57 148 DAPIN.
FT DOMAIN 280 602 NACHT.
FT REPEAT 704 727 LRR 1.
FT REPEAT 730 753 LRR 2.
FT REPEAT 780 803 LRR 3.
FT REPEAT 809 832 LRR 4.
FT REPEAT 836 863 LRR 5.
FT REPEAT 865 892 LRR 6.
FT REPEAT 893 916 LRR 7.
FT REPEAT 950 973 LRR 8.
FT REPEAT 979 1002 LRR 9.
FT REPEAT 1007 1034 LRR 10.
FT REPEAT 1036 1059 LRR 11.
FT REPEAT 1064 1092 LRR 12.
FT REPEAT 1121 1142 LRR 13.
FT NP_BIND 286 293 ATP (POTENTIAL).
SQ SEQUENCE 1200 AA; 134235 MW; 9A070D2A771B28FA CRC64;

Query Match 1.2%; Score 12; DB 1; Length 1200;
Best Local Similarity 100.0%; Pred. No. 0.0022;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 259 PERLLFIIDGFD 270
DB 354 PERLLFIIDGFD 365
|||||
|||||

RESULT 5
ID NAL7 HUMAN STANDARD; PRT; 980 AA.
AC Q8WY34;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 7 (PYRIN-containing APAF1-like
DE protein 3).
GN NALP7 OR PYPAF3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22162427; PubMed=12019269;
RA Wang L., Manji G.A., Grenier J.M., Al-Garawi A., Merriam S.,
RA Lora J.M., Geddes B.J., Briskin M., DiStefano P.S., Bertin J.;
RT "PYPAF7, a novel PYRIN-containing Apaf1-like protein that regulates
RT activation of NF-kappa B and caspase-1-dependent cytokine
RT processing.";
RL J. Biol. Chem. 277:29874-29880(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Tschopp J., Martinon F., Burns K.;
RT "NALPs: a novel protein family involved in inflammation.";
RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 9 leucine-rich (LRR) repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF464765; AAL69963.1; -.
CC EMBL; AY154462; AAO18158.1; -.
CC Genew; HGNC:22947; NALP7.
CC InterPro; IPR007091; LRR_RNinh.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF05729; NACHT; 1.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC PROSITE; PS50824; DAPIN; 1.
CC PROSITE; PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 93 DAPIN.
FT DOMAIN 172 491 NACHT.
FT REPEAT 614 638 LRR 1.
FT REPEAT 674 697 LRR 2.
FT REPEAT 760 784 LRR 3.
FT REPEAT 788 810 LRR 4.
FT REPEAT 817 840 LRR 5.
FT REPEAT 845 868 LRR 6.
FT REPEAT 874 897 LRR 7.
FT REPEAT 902 928 LRR 8.
FT REPEAT 933 957 LRR 9.
FT NP_BIND 178 185 ATP (POTENTIAL).
SQ SEQUENCE 980 AA; 111806 MW; 822AF2FD4338003D CRC64;

Query Match 0.9%; Score 9; DB 1; Length 980;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 YSFIHLSFQ 491
DB 464 YSFIHLSFQ 472
|||||
|||||

RESULT 6
ID NAL2 HUMAN STANDARD; PRT; 1062 AA.
AC Q9NX02; Q9BNV5; Q9H6G6; Q9HAV9; Q9NWK3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 2 (PYRIN-containing APAF1-like
DE protein 2) (Nucleotide-binding site protein 1).
GN NALP2 OR PYPAF2 OR NES1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF298547; AAG15253.1; ALT_INIT.
 DR EMBL; AF310106; AAG30289.1; -
 DR EMBL; AF464764; AAL69962.1; -
 DR EMBL; AK000517; BAA91223.1; -
 DR EMBL; AK025952; BAB15293.1; -
 DR EMBL; AK000784; BAA91377.1; ALT_INIT.
 DR EMBL; BC003592; AAH03592.1; -
 DR EMBL; BC001039; AAH01039.1; -
 DR HSSP; PI0775; 2BNH.
 DR Genew; HGNC:22948; NALP2.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
 KW Alternative splicing.
 FT DOMAIN 1 94 DAPIN.
 FT DOMAIN 207 526 NACHT.
 FT REPEAT 467 491 LRR 1.
 FT REPEAT 622 645 LRR 2.
 FT REPEAT 754 777 LRR 3.
 FT REPEAT 810 832 LRR 4.
 FT REPEAT 839 862 LRR 5.
 FT REPEAT 867 890 LRR 6.
 FT REPEAT 924 947 LRR 7.
 FT REPEAT 981 1005 LRR 8.
 FT REPEAT 1010 1033 LRR 9.
 FT NP_BIND 213 220 ATP (POTENTIAL).
 FT DOMAIN 518 523 POLY-GLU.
 FT VARSPLIC 133 154 Missing (in isoform 2).
 FT CONFLICT 1 1 M -> V (IN REF. 1).
 FT CONFLICT 35 35 L -> P (IN REF. 1).
 FT CONFLICT 304 304 I -> S (IN REF. 4; BAB15293).
 FT CONFLICT 364 364 R -> K (IN REF. 1).
 FT CONFLICT 980 980 MISSING (IN REF. 1).
 FT CONFLICT 1052 1052 A -> E (IN REF. 5).
 SQ SEQUENCE 1062 AA; 4DBB0F6E9C2BC8A7 CRC64;
 Query Match 0.9%; Score 9; DB 1; Length 1062;
 Best Local Similarity 100.0%; Pred. No. 2.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 483 YSFHLSFQ 491
 Db 499 YSFHLSFQ 507
 RESULT 7
 NA14 HUMAN STANDARD; PRT; 1093 AA.
 ID NA14 HUMAN
 AC Q86W24;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE NACHT-, LRR- and PYD-containing protein 14.
 GN NALP14.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22451042; PubMed=12563287;
 RA Tschopp J., Martinon F., Burns K.;
 ET "NALPs: a novel protein family involved in inflammation.";
 RL Nat. Rev. Mol. Cell Biol. 4:95-104(2003).

```

CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 11 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AV154469; AAO18165.1; -.
CC Genew: HGNC:22939; NALP14.
CC InterPro: IPR008938; ARM.
CC InterPro: IPR001611; LRR.
CC InterPro: IPR007091; LRR_RNinh.
CC InterPro: IPR003590; LRR_RNinh_sub.
CC InterPro: IPR007111; NACHT_NTPase.
CC InterPro: IPR004020; PAAD_DAPIN_dom.
CC Pfam: PF00560; LRR; 2.
CC Pfam: PF02758; PAAD_DAPIN; 1.
CC PRINTS: PR00019; LEURICHRPT.
CC SMART: SM00368; LRR_R1; 12.
CC PROSITE: PS50824; DAPIN; 1.
CC PROSITE: PS50837; NACHT; 1.
KW ATP-binding; Leucine-rich repeat; Repeat.
FT DOMAIN 1 97
FT DAPIN.
FT DOMAIN 177 499
FT NACHT.
FT REPEAT 525 550
FT LRR 1.
FT REPEAT 672 695
FT LRR 2.
FT REPEAT 733 756
FT LRR 3.
FT REPEAT 757 780
FT LRR 4.
FT REPEAT 785 807
FT LRR 5.
FT REPEAT 842 869
FT LRR 6.
FT REPEAT 871 894
FT LRR 7.
FT REPEAT 899 926
FT LRR 8.
FT REPEAT 956 979
FT LRR 9.
FT REPEAT 985 1008
FT LRR 10.
FT REPEAT 1013 1037
FT LRR 11.
SQ SEQUENCE 1093 AA; 124732 MW; 124EEACE22A11D6F CRC64;

Query Match 0.9%; Score 9; DB 1; Length 1093;
Best Local Similarity 100.0%; Pred. No. 2.5; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 711 LRHPNCKLQ 719
Db 697 LRHPNCKLQ 705
|||||
|

RESULT 8
ID NALL1 HUMAN STANDARD; PRT; 1473 AA.
AC Q9C000; Q9BZ28; Q9HAV8; Q9UFT4; Q9Y2B0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 42, Last annotation update)
DE NACHT-, LRR- and PYD-containing protein 2 (Death effector filament-
DE forming ced-4-like apoptosis protein) (Nucleotide-binding domain and
DE caspase recruitment domain) (Caspase recruitment domain protein 7).
GN NALP1 OR DEFCAP OR NAC OR CARD7 OR KIA0926.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21169419; PubMed=11270363;
RA Bertain J., DiStefano P.S.;
RT "The PYRIN domain: a novel motif found in apoptosis and inflammation
RT proteins."
RT Cell Death Differ. 7:1273-1274(2000).

```

```

RN RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=21148093; PubMed=11250163;
RA Martinon F., Hofmann K., Tschopp J.;
RT "The pyrin domain: a possible member of the death domain-fold family
RL Curr. Biol. 11:R118-R120(2001).
RN RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX TISSUE=Brythroleukemia;
RC MEDLINE=21153743; PubMed=11076957;
RA Hlaing T., Guo R.-F., Dilley K.A., Loussia J.M., Morrish T.A.,
RA Shi M.M., Vincenz C., Ward P.A.;
RT "Molecular cloning and characterization of DEFCAP-L and -S, two
RT isoforms of a novel member of the mammalian Ced-4 family of apoptosis
RT proteins."
RL J. Biol. Chem. 276:9230-9238(2001).
RN RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4), AND PROTEIN INTERACTION.
RC TISSUE=T-cell;
RX MEDLINE=21153744; PubMed=11113115;
RA Chu Z.-L., Pio F., Xie Z., Welsh K., Krajewska M., Krajewski S.,
RA Godzik A., Reed J.C.;
RT "A novel enhancer of the Apaf1 apoptosome involved in cytochrome
RT c-dependent caspase activation and apoptosis."
RL J. Biol. Chem. 276:9239-9245(2001).
RN RP SEQUENCE FROM N.A. (ISOFORM 2).
RX TISSUE=Brain;
RC MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K.-I., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RL for large proteins in vitro."
RN DNA Res. 6:63-70(1999).
RN RP SEQUENCE OF 282-1473 FROM N.A. (ISOFORM 1).
RC TISSUE=Uterus;
RA Koehler K., Beyer A., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Able to form cytoplasmic structures termed death
CC effector filaments. Enhances APAF1 and cytochrome c-dependent
CC activation of pro-caspase-9 and consecutive apoptosis. Seems to
CC bind ATP.
CC -!- SUBUNIT: Interacts strongly with caspase 2, weakly with caspase 9
CC and with APAF1 in a cytochrome c-inducible way leading to the
CC formation of an apoptosome. This interaction may be ATP-dependent.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=1; Synonyms=NAC beta, DEFCAP-L;
CC IsoId=Q9C000-1; Sequence=Displayed;
CC Name=2; Synonyms=NAC alpha, DEFCAP-S;
CC IsoId=Q9C000-2; Sequence=VSP_004327;
CC Name=3; Synonyms=NAC gamma;
CC IsoId=Q9C000-3; Sequence=VSP_004326, VSP_004327;
CC Name=4; Synonyms=NAC delta;
CC IsoId=Q9C000-4; Sequence=VSP_004326;
CC -!- TISSUE SPECIFICITY: Widely expressed. Isoforms 1 and 2 are
CC expressed in peripheral blood leukocytes, chronic myelogenous
CC leukemia cell line K-562, followed by thymus, spleen and heart.
CC Also detected in lung, placenta, small intestine, colon, kidney,
CC liver and muscle.
CC -!- SIMILARITY: Contains 1 DAPIN domain.
CC -!- SIMILARITY: Contains 1 NACHT domain.
CC -!- SIMILARITY: Contains 1 CARD domain.
CC -!- SIMILARITY: Contains 7 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF298548; AAG15254.1; -
CC EMBL; AF310105; AAG30288.1; -
CC EMBL; AF229059; AAK00748.1; -
CC EMBL; AF229060; AAK00749.1; -
CC EMBL; AF229061; AAK00750.1; -
CC EMBL; AF229062; AAK00751.1; -
CC EMBL; AB023143; BAA76770.1; -
CC EMBL; AL117470; CAB55945.1; -
CC EMBL; T17255; T17255.
CC HSP; P13489; IAA4Y.
CC MIM; 606636; -
CC GO; GO:0005622; C:intracellular; IC.
CC GO; GO:0016506; F:apoptosis activator activity; NAS.
CC GO; GO:0008566; F:caspase activator activity; NAS.
CC GO; GO:0019899; F:enzyme binding; IPI.
CC GO; GO:0006919; F:caspase activation; NAS.
CC GO; GO:0006917; F:induction of apoptosis; NAS.
CC InterPro; IPR001315; CARD.
CC InterPro; IPR000767; Disease_resist.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_Rninh.
CC InterPro; IPR007111; NACHT_NTPase.
CC InterPro; IPR004020; PAAD_DAPIN_dom.
CC Pfam; PF00560; LRR; 2.
CC Pfam; PF05729; NACHT; 1.
CC Pfam; PF02758; PAAD_DAPIN; 1.
CC PRINTS; PR00364; DISEASERIST.
CC PROSITE; PS02029; CARD; 1.
CC PROSITE; PS00824; DAPIN; 1.
CC PROSITE; PS00837; NACHT; 1.
CC Apoptosis; ATP-binding; Leucine-rich repeat; Repeat;
KW Alternative splicing.
FT DOMAIN 1 92
FT DAPIN.
FT DOMAIN 328 637
FT NACHT.
FT REPEAT 704 725
FT LRR 1.
FT REPEAT 807 830
FT LRR 2.
FT REPEAT 864 887
FT LRR 3.
FT REPEAT 921 944
FT LRR 4.
FT REPEAT 950 973
FT LRR 5.
FT REPEAT 1199 1215
FT LRR 6.
FT REPEAT 1216 1236
FT LRR 7.
FT DOMAIN 1374 1463
FT NP BIND 334 341
FT VARSPLIC 958 987
FT VARSPLIC 1262 1305
FT VARSPLIC 1262 1305
FT MUTAGEN 340 340
FT MUTAGEN 340 340
FT CONFLICT 155 155
FT CONFLICT 246 246
FT CONFLICT 782 782
FT CONFLICT 878 878
FT CONFLICT 995 995
FT CONFLICT 1119 1119
FT CONFLICT 1184 1184
FT CONFLICT 1241 1241
FT CONFLICT 1366 1366
FT CONFLICT 1473 AA; 165865 NW; 438F0DCE45C2562D CRC64;
SQ SEQUENCE 1473 AA; 165865 NW; 438F0DCE45C2562D CRC64;
Query Match 0.9%; Score 9; DB 1; Length 1473;
Best Local Similarity 100.0%; Pred.No. 3.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 189 QGAGIGKGS 197
DB 333 QGAGIGKGS 341

RESULT 9
UCN1 RAT STANDARD; PRT; 122 AA.
ID UCN1 RAT STANDARD; PRT; 122 AA.
AC P55030;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Urocortin precursor (Corticotensin).
GN UCN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=96069764; PubMed=7477349;
RA Vaughan J.M., Donaldson C.J., Bittencourt J., Perrin M.H., Lewis K.A.,
Sutton S.W., Chan R., Turnbull A., Lovejoy D., Rivier C., Rivier J.E.,
Sawchenko P., Vale W.W.;
RA "Urocortin, a mammalian neuropeptide related to fish urotensin I and
RT to corticotropin-releasing factor.";
RL Nature 378:287-292 (1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Lewis;
RA Park J.H., Lee Y.J., Kim K.L.;
RT "Detection of rat urocortin in lymphoid tissues: implications for the
functional assessment of urocortin as a novel neuro-immunomodulatory
peptide.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Acts in vitro to stimulate the secretion of
adrenocorticotrophic hormone (ACTH). Binds with high affinity to
CC CRF Receptor types 1, 2-alpha, and 2-beta.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the sauvagine/corticotropin-releasing
factor/urotensin I family.
CC -----
CC EMBL; U33935; AAA87566.1; -
CC EMBL; AF093623; AAF63153.1; -
CC PIR; S60262; S60262.
CC InterPro; IPR000187; corticoliberin.
CC InterPro; IPR003620; Urocortin_CRF.
CC Pfam; PF00473; CRF; 1.
CC PRINTS; PR01612; CRFFAMILY.
CC SMART; SM00039; CRF; 1.
CC PROSITE; PS00511; CRF; 1.
CC Hormone; Amidation; Cleavage on pair of basic residues; Signal.
KW SIGNAL 1 25
FT PROPEP 26 80
FT PEPTIDE 81 120
FT MOD_RES 120 120
FT AMIDATION (G-121 PROVIDE AMIDE GROUP) (BY
FT SIMILARITY).
SQ SEQUENCE 122 AA; 13711 MW; 9F0AF834CBFCET74 CRC64;
Query Match 0.8%; Score 8; DB 1; Length 122;
Best Local Similarity 100.0%; Pred.No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 LLVQLRPE 667
DB 15 LLVQLRPE 22
RESULT 10

```

RINI_PIG
ID RINI_PIG STANDARD; PRT; 456 AA.
AC P10775;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease inhibitor.
DE RNH OR RI.
GN Sus scrofa (Pig).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE
RC TISSUE=Liver;
RX MEDLINE=89118268; PubMed=3219361;
RA Hofsteenge J., Kieffer B., Matthies R., Hemmings B.A., Stone S.R.;
RT "Amino acid sequence of the ribonuclease inhibitor from porcine liver
RL reveals the presence of leucine-rich repeats.";
RL Biochemistry 27:8537-8544(1988).
RN [2]
RP SEQUENCE OF 82-456 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=91104783; PubMed=2271559;
RA Vicentini A.M., Kieffer B., Matthies R., Meyhack B., Hemmings B.A.,
RA Stone S.R., Hofsteenge J.;
RT "Protein chemical and kinetic characterization of recombinant porcine
RL ribonuclease inhibitor expressed in Saccharomyces cerevisiae.";
RL Biochemistry 29:8827-8834(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=94088748; PubMed=8264799;
RA Kobe B., Deisenhofer J.;
RT "Crystal structure of porcine ribonuclease inhibitor, a protein with
RL leucine-rich repeats.";
RL Nature 366:751-756(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=95183144; PubMed=7877692;
RA Kobe B., Deisenhofer J.;
RT "A structural basis of the interactions between leucine-rich repeats
RL and protein ligands.";
RL Nature 374:183-186(1995).
CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiotensin. May also
CC function in the modulation of cellular activities.
CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58700; AAA63448.1; -.
DR F01; A31857; A31857.
DR PDB; 2NH; 12-MAR-97.
DR PDB; 1DFJ; 11-JAN-97.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_Rninh.
DR InterPro; IPR003590; LRR_Rninh_sub.
DR SMART; SM00368; LRR_R1; 2.
DR Pfam; PF00560; LRR; 5.
KW Acetylation; Repeat; Leucine-rich repeat; 3D-structure.
FT REPEAT 15 43 LRR A1.
FT REPEAT 44 71 LRR B1.
FT REPEAT 72 100 LRR A2.
FT REPEAT 101 128 LRR B2.
FT REPEAT 129 157 LRR A3.
FT REPEAT 158 185 LRR B3.
FT

```

```

FT REPEAT 214 214 LRR A4.
FT REPEAT 215 242 LRR B4.
FT REPEAT 243 271 LRR A5.
FT REPEAT 272 299 LRR B5.
FT REPEAT 300 328 LRR A6.
FT REPEAT 329 356 LRR B6.
FT REPEAT 357 385 LRR A7.
FT REPEAT 386 413 LRR B7.
FT REPEAT 414 442 LRR A8.
FT MOD RES 1 1 ACETYLATION.
FT STRAND 2 2
FT STRAND 5 6
FT HELIX 12 22
FT TURN 23 24
FT STRAND 27 31
FT TURN 32 32
FT HELIX 37 47
FT TURN 48 49
FT TURN 51 52
FT STRAND 55 57
FT HELIX 64 75
FT TURN 78 79
FT STRAND 84 86
FT TURN 88 89
FT HELIX 94 98
FT TURN 100 106
FT STRAND 108 109
FT HELIX 112 114
FT TURN 121 133
FT TURN 135 136
FT STRAND 141 143
FT TURN 145 146
FT STRAND 150 150
FT STRAND 169 171
FT TURN 173 174
FT STRAND 176 176
FT HELIX 178 180
FT TURN 191 191
FT STRAND 198 200
FT TURN 202 203
FT HELIX 208 220
FT TURN 222 223
FT STRAND 226 228
FT HELIX 235 246
FT TURN 247 247
FT TURN 249 250
FT STRAND 255 257
FT TURN 259 260
FT HELIX 265 277
FT STRAND 283 285
FT TURN 287 288
FT HELIX 292 303
FT TURN 304 304
FT TURN 306 307
FT STRAND 312 314
FT TURN 316 317
FT STRAND 321 321
FT HELIX 322 324
FT STRAND 325 334
FT STRAND 340 342
FT STRAND 347 347
FT HELIX 349 359
FT TURN 360 361
FT STRAND 369 371
FT TURN 373 374
FT HELIX 379 391
FT STRAND 397 399
FT TURN 401 402
FT HELIX 407 417
FT TURN 418 418
FT STRAND 426 428
FT TURN 430 431

```

```
FT HELIX 436 448
FT STRAND 453 456
SQ SEQUENCE 456 AA; 49023 MW; 01DA0A529CDC763E CRC64;

Query Match 0.8%; Score 8; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SRELDLS 866
DB 394 SRELDLS 401

RESULT 11
RINI_RAT
ID RINI_RAT STANDARD; PRT; 456 AA.
AC P29315;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ribonuclease inhibitor.
DE RNH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=92162755; PubMed=1536887;
RA Kawanomoto M., Motojima K., Sasaki M., Hattori H., Goto S.;
RT "cDNA cloning and sequence of rat ribonuclease inhibitor, and tissue
distribution of the mRNA.";
RL Biochim. Biophys. Acta 1129:335-338(1992).
CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
function in the modulation of cellular activities.
CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Brain, heart, lung, liver, spleen, testes and
kidney; highest in the lung and lowest in the heart.
CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X62528; CAA4388.1; --
CC F1R; S20597; S20597.
CC HSSP; P10775; 2BNH.
CC InterPro; IPR001611; LRR.
CC InterPro; IPR007091; LRR_Rninh.
CC Pfam; PF00560; LRR_4.
CC SMART; SM00368; LRR_R1; 1.
KW Repeat; Leucine-rich repeat.
FT REPEAT 15 43 LRR A1.
FT REPEAT 44 71 LRR B1.
FT REPEAT 72 100 LRR A2.
FT REPEAT 101 128 LRR B2.
FT REPEAT 129 157 LRR A3.
FT REPEAT 158 185 LRR B3.
FT REPEAT 186 214 LRR A4.
FT REPEAT 215 242 LRR A5.
FT REPEAT 243 271 LRR B5.
FT REPEAT 272 299 LRR B6.
FT REPEAT 300 328 LRR A6.
FT REPEAT 329 356 LRR B7.
FT REPEAT 357 385 LRR A7.
FT REPEAT 386 413 LRR B7.

FT REPEAT 414 442 LRR A8.
SQ SEQUENCE 456 AA; 49905 MW; 8518E5B1F09E5998 CRC64;

Query Match 0.8%; Score 8; DB 1; Length 456;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 830 CRLRTLWL 837
DB 251 CRLRTLWL 258

RESULT 12
RINI_HUMAN
ID RINI_HUMAN STANDARD; PRT; 460 AA.
AC P13459;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental ribonuclease inhibitor (ribonuclease/angiogenin inhibitor)
DE (RAI) (RNase inhibitor) (RI).
DE RNH OR PRI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9118269; PubMed=3219362;
RA Lee F.S., Fox E.A., Zhou H.-M., Strydom D.J., Vallee B.L.;
RT "Primary structure of human placental ribonuclease inhibitor.";
RL Biochemistry 27:8545-8553(1988).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 443-461.
RX MEDLINE=9210799; PubMed=3243277;
RA Schneider R., Schneider-Scherzer E., Thurnher M., Auer B.,
RA Schweiger M.;
RT "The primary structure of human ribonuclease/angiogenin inhibitor
(RAI) discloses a novel highly diversified protein superfamily with a
common repetitive module.";
RL EMBO J. 7:4151-4156(1988).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Poustka A., Wellenreuther R., Mewes H.-W., Weil B., Wiemann S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, Kidney, Lymph, and Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF COMPLEX WITH ANGIOGENIN.
RX MEDLINE=97459904; PubMed=9311977;
```


RA Papageorgiou A.C., Shapiro R., Acharya K.R.;
 RT "Molecular recognition of human angiogenin by placental ribonuclease
 RL inhibitor: -- an X-ray crystallographic study at 2.0-A resolution.";
 EMBO J. 16:5162-5177(1997).
 CC -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
 CC function in the modulation of cellular activities.
 CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; M22414; AAA59130.1; -;
 CC EMBL; X13973; CAA32151.1; -;
 CC EMBL; N36717; AAA60249.1; -;
 CC EMBL; AL161967; CAB82310.1; -;
 CC EMBL; BC003075; AAH03075.1; -;
 CC EMBL; BC011500; AAH11500.1; -;
 CC EMBL; BC014629; AAH14629.1; -;
 CC EMBL; BC047730; AAH47730.1; -;
 CC PIR; A31858; A31858.
 CC PDB; 1A4Y; 14-OCT-98.
 CC Genew; HGNC:10074; RNH.
 CC MIM; 173320; -;
 CC GO; GO:0008428; Ribonuclease inhibitor activity; TAS.
 CC GO; GO:0006401; P:RNA catabolism; TAS.
 CC InterPro; IPR001611; LRR.
 CC InterPro; IPR007091; LRR_RNinh.
 CC InterPro; IPR003590; LRR_RNinh_sub.
 CC Pfam; PF00560; LRR; 3.
 CC SMART; SM00368; LRR_R1; 1.
 KW Repeat; Leucine-rich repeat; 3D-structure; Polymorphism.
 FT INIT_MET 0 0
 FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
 FT REPEAT 19 47 LRR A1.
 FT REPEAT 48 75 LRR B1.
 FT REPEAT 76 104 LRR A2.
 FT REPEAT 105 132 LRR B2.
 FT REPEAT 133 161 LRR A3.
 FT REPEAT 162 189 LRR B3.
 FT REPEAT 190 218 LRR A4.
 FT REPEAT 219 246 LRR B4.
 FT REPEAT 247 275 LRR A5.
 FT REPEAT 276 303 LRR B5.
 FT REPEAT 304 332 LRR A6.
 FT REPEAT 333 360 LRR B6.
 FT REPEAT 361 389 LRR A7.
 FT REPEAT 390 417 LRR B7.
 FT REPEAT 418 446 LRR A8.
 FT VARIANT 169 169 P -> L (in dbSNP:17585).
 FT
 FT CONFLICT 422 423 /FTId-VAR_014726.
 FT STRAND 2 10 RQ -> SE (IN REF. 2).
 FT HELIX 16 22
 FT TURN 23 28
 FT STRAND 31 35
 FT TURN 41 43
 FT HELIX 44 52
 FT TURN 53 53
 FT STRAND 59 61
 FT TURN 63 64
 FT HELIX 68 76
 FT TURN 77 79
 FT TURN 82 83
 FT STRAND 88 90
 FT TURN 92 93
 FT HELIX 98 100

FT HELIX 101 110
 FT TURN 112 113
 FT STRAND 116 118
 FT HELIX 125 136
 FT TURN 137 137
 FT TURN 139 140
 FT STRAND 145 147
 FT TURN 149 150
 FT HELIX 155 157
 FT HELIX 158 167
 FT TURN 169 170
 FT STRAND 173 175
 FT HELIX 182 195
 FT STRAND 202 204
 FT TURN 206 207
 FT TURN 213 213
 FT HELIX 214 224
 FT TURN 226 227
 FT STRAND 230 232
 FT HELIX 239 250
 FT TURN 251 251
 FT TURN 253 254
 FT STRAND 259 261
 FT TURN 263 264
 FT HELIX 269 279
 FT TURN 280 281
 FT TURN 283 284
 FT STRAND 287 289
 FT TURN 291 292
 FT HELIX 296 307
 FT TURN 308 308
 FT TURN 310 311
 FT STRAND 316 318
 FT TURN 320 321
 FT STRAND 325 325
 FT STRAND 326 328
 FT HELIX 329 338
 FT STRAND 344 346
 FT STRAND 351 351
 FT HELIX 353 364
 FT TURN 365 365
 FT TURN 367 368
 FT STRAND 373 375
 FT TURN 377 378
 FT TURN 383 385
 FT HELIX 386 395
 FT STRAND 401 403
 FT TURN 405 406
 FT HELIX 411 422
 FT TURN 424 425
 FT STRAND 430 432
 FT TURN 434 435
 FT HELIX 440 452
 FT TURN 454 455
 FT STRAND 457 460
 SQ SEQUENCE 460 AA; 49842 MW; C3D6668E2F2BFB6F CRC64;

Query Match 0.8%; Score 8; DB 1; Length 460;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 859 SIRELDLS 866
 Db 398 SIRELDLS 405

RESULT 13

RINI_PANTR STANDARD; PRT; 460 AA.
 ID RINI_PANTR
 AC Q8HZE9;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Placental ribonuclease inhibitor (Ribonuclease/angiogenin inhibitor)
 DE (RAI) (RNase inhibitor) (RI).
 GN RNH.
 OS Pan troglodytes (Chimpanzee).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Zhang J., Webb D.M., Podlaha O.;
 RA "Accelerated protein evolution and origins of human-specific features:
 RT FOX2 as an example.";
 RT Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RL -!- FUNCTION: Inhibitor of pancreatic RNase and angiogenin. May also
 CC function in the modulation of cellular activities (By similarity).
 CC -!- SUBUNIT: Forms a tight one-to-one complex with the RNase (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- SIMILARITY: Contains 15 leucine-rich (LRR) repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF539549; AAN10133.1; -.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR006553; LRR_CYS_sub.
 DR InterPro; IPR007091; LRR_Rninh.
 DR InterPro; IPR003590; LRR_Rninh_sub.
 DR Pfam; PF00560; LRR; 3.
 DR PRINTS; PR00019; LEURICHRPT.
 DR SMART; SM00367; LRR_CC; 5.
 DR SMART; SM00368; LRR_RI; 13.
 DR Repeat; Leucine-rich repeat.
 KW INIT MET 0 BY SIMILARITY.
 FT DOMAIN 1 10 2 X 5 AA TANDEM REPEATS OF S-L-D-I-Q.
 FT REPEAT 19 47 LRR A1.
 FT REPEAT 48 75 LRR B1.
 FT REPEAT 76 104 LRR A2.
 FT REPEAT 105 132 LRR B2.
 FT REPEAT 133 161 LRR A3.
 FT REPEAT 162 189 LRR B3.
 FT REPEAT 190 218 LRR A4.
 FT REPEAT 219 246 LRR B4.
 FT REPEAT 247 275 LRR A5.
 FT REPEAT 276 303 LRR B5.
 FT REPEAT 304 332 LRR A6.
 FT REPEAT 333 360 LRR B6.
 FT REPEAT 361 389 LRR A7.
 FT REPEAT 390 417 LRR B7.
 FT REPEAT 418 446 LRR A8.
 SQ SEQUENCE 460 AA; 49798 MW; C11F15EE97DBB4F CRC64;
 Query Match 0.8%; Score 8; DB 1; Length 460;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 859 SIRELDLS 866
 DB 398 SIRELDLS 405
 RESULT 14
 VC41_BPT4 STANDARD; PRT; 475 AA.
 AC P04530;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Primase-helicase (Protein Gp41).
 GN 41.
 OS Bacteriophage T4.
 OS Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
 OC T4-like viruses.
 OX NCBI_TaxID=10685;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA MEDLINE=89340565; PubMed=2668290;
 RA Nakanishi M., Alberts B.;
 RA Unpublished results, cited by:
 RL Hinton D.M.;
 RL J. Biol. Chem. 264:14440-14446(1989).
 RL [2]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=22514363; PubMed=12626685;
 RX Miller E.S., Kutter E., Mosig G., Arisaka F., Kunisawa T., Ruger W.;
 RA "Bacteriophage T4 genome.";
 RT Microbiol. Mol. Biol. Rev. 67:86-156(2003).
 RN [3]
 RN CHARACTERIZATION.
 RP MEDLINE=94149695; PubMed=8107085;
 RP Young M.C., Schlutz D.E., Ring D., von Hippel P.H.;
 RA "Kinetic parameters of the translocation of bacteriophage T4 gene 41
 RT protein helicase on single-stranded DNA.";
 RL J. Mol. Biol. 235:1447-1458(1994).
 CC -!- FUNCTION: Essential replication protein, part of the primase-
 CC helicase required for lagging strand DNA synthesis. It acts
 CC processively. It forms a protein complex with the gene 59 protein
 CC that partly replaces the dda protein helicase function. Act as
 CC single-stranded ATP-dependent DNA helicase.
 CC -!- MISCELLANEOUS: Interacts with the gene 61 protein to form the T4
 CC primosome.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; K03113; AAA32553.1; -.
 DR EMBL; AF158101; AAD42466.1; -.
 DR PIR; A04308; Z4BPT4.
 DR InterPro; IPR007694; DnaB_C.
 DR Pfam; PF03796; DnaB_C; 1.
 DR Helicase; DNA replication; ATP-binding; DNA-binding.
 KW NP BIND 197 204 ATP (POTENTIAL).
 FT SEQUENCE 475 AA; 53601 MW; A544B1F9CFC90B4C CRC64;
 SQ
 Query Match 0.8%; Score 8; DB 1; Length 475;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 433 GLCSLAAD 440
 DB 206 GLCSLAAD 213
 RESULT 15
 C180_MOUSE
 ID_C180_MOUSE STANDARD; PRT; 661 AA.
 AC Q62192;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE CD180 antigen precursor (Lymphocyte antigen 78) (Radioprotective 105
 DE kDa protein).
 DE LY78 OR RP105.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A., AND SEQUENCE OF 21-43.
STRAIN=BALE/C; TISSUE=B-cell lymphoma;
RA MEDLINE=95204928; PubMed=7897216;
RA Miyake K., Yamashita Y., Ogata M., Sudo T., Kimoto M.;
RT "RP105, a novel B cell surface molecule implicated in B cell
activation, is a member of the leucine-rich repeat protein family."
RL J. Immunol. 154:3333-3340(1995).
[2]
INTERACTION WITH MD-1.
RA MEDLINE=98349386; PubMed=9686597;
RA Miyake K., Shimazu R., Kondo J., Niki T., Akashi S., Ogata H.,
RA Yamashita Y., Miura Y., Kimoto M.;
RT "Mouse MD-1, a molecule that is physically associated with RP105 and
positively regulates its expression."
RL J. Immunol. 161:1348-1353(1998).
[3]
FUNCTION.
RX MEDLINE=20341616; PubMed=10880523;
RA Ogata H., Su I., Miyake K., Nagai Y., Akashi S., Mecklenbrauer I.,
RA Rajewsky K., Kimoto M., Tarakhovskiy A.;
RT "The Toll-like receptor protein RP105 regulates lipopolysaccharide
signaling in B cells."
RL J. Exp. Med. 192:23-29(2000).
CC -!- FUNCTION: May cooperate with MD-1 and TLR4 to mediate the innate
immune response to bacterial lipopolysaccharide (LPS) in B cells.
CC Leads to NF-kappa-B activation. Also involved in the life/death
decision of B cells.
CC -!- SUBUNIT: Binds to MD-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein; plasma membrane.
CC -!- TISSUE SPECIFICITY: B lymphocytes and spleen. Not detected in
thymus, kidney, muscle, heart, brain or liver.
CC -!- SIMILARITY: Belongs to the Toll-like receptor family.
CC -!- SIMILARITY: Contains 17 leucine-rich (LRR) repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; D37797; BAA07043.1; --
DR PIR; I56258; I56258.
DR MGI; MGI:1194924; Ly78.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR000483; LRR_Cterm.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF01463; LRRCT; 1.
DR SMART; SM00082; LRRCT; 1.
KW Receptor; Immune response; Inflammatory response; Signal; Antigen;
KW Transmembrane; Repeat; Leucine-rich repeat; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 661 CD180 ANTIGEN.
FT DOMAIN 21 626 EXTRACELLULAR.
FT TRANSMEM 627 650 POTENTIAL.
FT DOMAIN 651 661 CYTOPLASMIC.
FT REPEAT 51 75 LRR 1.
FT REPEAT 76 99 LRR 2.
FT REPEAT 101 123 LRR 3.
FT REPEAT 148 172 LRR 4.
FT REPEAT 174 195 LRR 5.
FT REPEAT 199 221 LRR 6.
FT REPEAT 273 296 LRR 7.
FT REPEAT 297 320 LRR 8.
FT REPEAT 322 346 LRR 9.
FT REPEAT 369 392 LRR 10.
FT REPEAT 395 418 LRR 11.
FT REPEAT 420 443 LRR 12.
FT REPEAT 444 467 LRR 13.
FT REPEAT 468 490 LRR 14.

FT REPEAT 495 518
FT REPEAT 520 540
FT REPEAT 541 565
FT CARBOHYD 34 34 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 70 70 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 78 78 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 201 201 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 394 394 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 22 22 T -> D (IN REF. 1; AA SEQUENCE).
FT CONFLICT 24 24 S -> N (IN REF. 1; AA SEQUENCE).
FT CONFLICT 28 28 C -> L (IN REF. 1; AA SEQUENCE).
SQ SEQUENCE 661 AA; 74266 MW; 9A04386369803C9E CRC64;

Query Match 0.8%; Score 8; DB 1; Length 661;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 916 NURELDS 923
DB 371 NURELDS 378
|||||

Search completed: July 30, 2004, 14:04:15
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 30, 2004, 13:56:57 ; Search time 47 Seconds
(without alignments)
6948.114 Million cell updates/sec

Title: US-10-781-294-24
Perfect score: 1035
Sequence: 1 MURTAGRGLCLSTYLEL.....MTHSLAALRVTKPYLDIGC 1035

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0
Searched: 1017041 seqs, 315518202 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25:*
1: sp archea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp virus:*
16: sp bacteriaph:*
17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	740	71.5	892	Q8NEU4	Q8neu4 homo sapien
2	172	16.6	287	Q9BY26	Q9by26 homo sapien
3	12	1.2	1200	Q86W29	Q86w29 homo sapien
4	11	1.1	1043	Q86W25	Q86w25 homo sapien
5	10	1.0	501	Q8UZH1	Q8uzh1 cercopithec
6	9	0.9	509	Q9H5Z8	Q9h5z8 homo sapien
7	9	0.9	713	Q95LZ7	Q95lz7 macaca fasc
8	9	0.9	846	Q8IXT0	Q8ixt0 homo sapien
9	9	0.9	1065	Q9LGI5	Q9lgis oryza sativ
10	9	0.9	1093	Q86W24	Q86w24 homo sapien
11	9	0.9	1375	Q86UB5	Q86ub5 homo sapien
12	8	0.8	43	Q8KA86	Q8ka86 chlorobium
13	8	0.8	87	Q8ZB55	Q8zb55 yersinia pe
14	8	0.8	93	Q8FR50	Q8fr50 corynebacte
15	8	0.8	123	Q92LP7	Q92lp7 helicobacte
16	8	0.8	123	O25306	O25306 helicobacte

17	8	0.8	136	3	Q872H2	Q872h2 neurospora
18	8	0.8	136	16	Q70V25	Q7uv25 rhodospirell
19	8	0.8	143	2	Q50035	Q50035 mycobacteri
20	8	0.8	193	2	Q849M8	Q849m8 streptomyce
21	8	0.8	217	10	Q84S45	Q84s45 oryza sativ
22	8	0.8	237	4	Q9UOE6	Q9ue6 homo sapien
23	8	0.8	250	10	Q94L59	Q94l59 rhamnus cal
24	8	0.8	250	10	Q94L58	Q94l58 rhamnus cal
25	8	0.8	274	2	Q9F5H6	Q9f5h6 agrobacteri
26	8	0.8	274	2	Q9R725	Q9r725 agrobacteri
27	8	0.8	280	4	Q9UOE5	Q9ue5 homo sapien
28	8	0.8	284	11	Q9CV31	Q9cv31 mus musculu
29	8	0.8	300	16	Q92L81	Q92l81 rhizobium m
30	8	0.8	313	16	Q93LC6	Q93lc6 rhizobium m
31	8	0.8	317	4	Q9Y6K4	Q9y6k4 homo sapien
32	8	0.8	327	10	Q38695	Q38695 actinidia d
33	8	0.8	336	6	Q7YRA5	Q7yra5 bos taurus
34	8	0.8	338	16	Q88H07	Q88h07 pseudomonas
35	8	0.8	360	4	Q15435	Q15435 homo sapien
36	8	0.8	361	11	Q9Z105	Q9z105 mus musculu
37	8	0.8	364	16	Q8EX07	Q8ex07 mycoplasma
38	8	0.8	419	16	Q7UC70	Q7uc70 shigella fl
39	8	0.8	427	16	Q83ML0	Q83ml0 shigella fl
40	8	0.8	441	16	Q8PQP6	Q8qp6 xanthomonas
41	8	0.8	447	4	Q96FD7	Q96fd7 homo sapien
42	8	0.8	456	11	Q924P4	Q924p4 mus musculu
43	8	0.8	456	11	Q91V17	Q91v17 mus musculu
44	8	0.8	458	9	Q7Y596	Q7y596 bacterioph
45	8	0.8	461	4	Q81ZK8	Q81zk8 homo sapien

ALIGNMENTS

RESULT 1
Q8NEU4 PRELIMINARY; PRT; 892 AA.
AC Q8NEU4; (TREMREL. 22, Created)
DT 01-OCT-2002 (TREMREL. 22, Last sequence update)
DT 01-OCT-2002 (TREMREL. 22, Last sequence update)
DE Monarch-1 splice form IV.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams K.L., Linhoff M.W., Ting J.P.Y.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV116207; AAM75145.1; -;
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR RNinh.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 1.
DR Pfam; PF05729; NACHT; 1.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PROSITE; PS50824; DAPIN; 1.
DR PROSITE; PS50837; NACHT; 1.
SQ SEQUENCE 892 AA; 101733 MW; AE703D8DF341C2AC CRC64;

Query Match 71.5%; Score 740; DB 4; Length 892;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 740; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 97 DPQETVYVVRKFRIMEDRNARLGCYNLSHRYTLLLVKHSNPMVOVQQLDTRGRH 156
DB 124 DPQETVYVVRKFRIMEDRNARLGCYNLSHRYTLLLVKHSNPMVOVQQLDTRGRH 183
QY 157 ARTVGHQASPIKLTIFEPDEERPPPTVTVVQGAAGIGKSLAHKVMLDWADGKLFQGR 216
DB 184 ARTVGHQASPIKLTIFEPDEERPPPTVTVVQGAAGIGKSLAHKVMLDWADGKLFQGR 243

Nat. Rev. Mol. Cell Biol. 4:95-104 (2003).
 EMBL; AY154468; AA018164.1; -
 GO; GO:0005351; P:sugar porter activity; IEA.
 GO; GO:0009401; P:phosphoenolpyruvate-dependent sugar phospho. .; IEA.
 InterPro; IPR002114; Hpr_Serp_S.
 InterPro; IPR001611; LRR_RNinh.
 InterPro; IPR007091; LRR_RNinh.
 InterPro; IPR003590; LRR_RNinh sub.
 InterPro; IPR007111; NACHT_NTPase.
 InterPro; IPR004020; PAAD_DAPIN_dom.
 Pfam; PF00560; LRR_2.
 Pfam; PF02758; PAAD_DAPIN; 1.
 SMART; SM00368; LRR_RI; 9.
 PROSITE; PS0824; DAPIN; 1.
 PROSITE; PS0837; NACHT; 1.
 PROSITE; PS0589; PPS_HPR_SER; 1.
 SEQUENCE 1043 AA; 11910 MW; C321FBBG65206DFF CRC64;
 Query Match 1.1%; Score 11; DB 4; Length 1043;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 487 HLSFOFFFAAM 497
 DB 536 HLSFOFFFAAM 546
 RESULT 5
 QBUZHI PRELIMINARY; PRT; 501 AA.
 ID QBUZHI1
 AC QBUZHI1
 DT 01-MAR-2002 (TRENBLrel. 20, Created)
 DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
 DE BRRF2.
 OS Cercopithecine herpesvirus 15.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gamaherpesvirinae; Lymphocryptovirus.
 OX NCBI_TaxID=104228;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=97048062; PubMed=8892903;
 RA Franken M., Devergne O., Rosenzweig M., Annis B., Kieff E., Wang F.;
 RT "Comparative analysis identifies conserved tumor necrosis factor
 receptor-associated factor 3 binding sites in the human and simian
 Epstein-Barr virus oncogene LMPI.";
 RL J. Virol. 70:7819-7826 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=994124110; PubMed=10482645;
 RA Rivallier P., Quink C., Wang F.;
 RT "Strong selective pressure for evolution of an Epstein-Barr virus
 LMP2B homologue in the rhesus lymphocryptovirus.";
 RL J. Virol. 73:8867-8872 (1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=20304984; PubMed=10846073;
 RA Jiang H., Cho Y.-G., Wang F.;
 RT "Structural, functional, and genetic comparisons of Epstein-Barr virus
 nuclear antigen 3A, 3B, and 3C homologues encoded by the rhesus
 lymphocryptovirus.";
 RL J. Virol. 74:5921-5932 (2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=20440633; PubMed=10970361;
 RA Rao P., Jiang H., Wang F.;
 RT "Cloning of the rhesus lymphocryptovirus viral capsid antigen and
 Epstein-Barr virus-encoded small RNA homologues and use in diagnosis
 of acute and persistent infections.";

J. Clin. Microbiol. 38:3219-3225 (2000).
 [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RX MEDLINE=21602573; PubMed=11739708;
 RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
 RT "Complete Nucleotide Sequence of the Rhesus Lymphocryptovirus: Genetic
 Validation for an Epstein-Barr Virus Animal Model.";
 RL J. Virol. 76:421-426 (2002).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Mughaddam A., Koch J., Annis B., Wang F.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Mughaddam A., Annis B., Wang F.;
 RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Rivallier P., Quink C., Wang F.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Jiang H., Wang F.;
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [10]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Rao P.V., Jiang H., Wang F.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [11]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LCL8664;
 RA Rivallier P., Jiang H., Cho Y.-G., Quink C., Wang F.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY037858; AAK95439.1; -
 DR InterPro; IPR008550; DUF832.
 DR Pfam; PF05734; DUF832; 1.
 SQ SEQUENCE 501 AA; 53663 MW; 23BE56494C3FD3D6 CRC64;
 Query Match 1.0%; Score 10; DB 12; Length 501;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 946 LQKLWLDSCG 955
 DB 122 LQKLWLDSCG 131
 RESULT 6
 Q9H5Z8 PRELIMINARY; PRT; 509 AA.
 ID Q9H5Z8
 AC Q9H5Z8
 DT 01-MAR-2001 (TRENBLrel. 16, Created)
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
 DE Hypothetical protein FLJ22740.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawakami T., Neguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,
 RA Yamada K., Fujii Y., Ozaki K., Hiro M., Ohmori Y., Ota T., Suzuki Y.,
 RA Obayashi M., Nishi T., Shibahara T., Tanaka T., Nakamura Y.,
 RA Isegaki T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";

RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK026393; BAB15469.1; -.
 DR HSSP; P10775; 2BNH.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_RNinh.
 DR Pfam; PF00560; LRR; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 509 AA; EB7535AF69817F5B CRC64;

Query Match 0.9%; Score 9; DB 4; Length 509;
 Best Local Similarity 100.0%; Pred. No. 7.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 763 LLCEGLRHP 771
 |||||
 DB 207 LLCEGLRHP 215

RESULT 7
 Q95LZ7 PRELIMINARY; PRT; 713 AA.
 AC Q95LZ7
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Hashimoto K., Osada N., Hida M., Kusuda J., Tanuma R., Hirai M.,
 RA Terao K., Sugano S.;
 RT "Isolation of novel full-length cDNA clones from macaque testis cDNA
 RT libraries."
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB071044; BAB64437.1; -.
 DR InterPro; IPR007091; LRR_RNinh.
 DR InterPro; IPR007091; LRR_NTase.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS0837; NACHT; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 713 AA; 80258 MW; 6F214C9B773F54DC CRC64;

Query Match 0.9%; Score 9; DB 6; Length 713;
 Best Local Similarity 100.0%; Pred. No. 9.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 YSFIHLSFQ 491
 |||||
 DB 239 YSFIHLSFQ 247

RESULT 8
 Q8IXT0 PRELIMINARY; PRT; 846 AA.
 AC Q8IXT0
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to NALP2 protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC039269; AAH39269.1; -.
 DR InterPro; IPR007111; NACHT_NTPase.
 DR InterPro; IPR004020; PAAD_DAPIN_dom.
 DR Pfam; PF05729; NACHT; 1.
 DR PROSITE; PS0824; DAPIN; 1.
 DR PROSITE; PS0837; NACHT; 1.
 SQ SEQUENCE 846 AA; 96368 MW; 46BE824550E39F7 CRC64;

Query Match 0.9%; Score 9; DB 4; Length 846;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 483 YSFIHLSFQ 491
 |||||
 DB 499 YSFIHLSFQ 507

RESULT 9
 Q9LGI5 PRELIMINARY; PRT; 1065 AA.
 AC Q9LGI5
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 GN P009G03.21 OR P0030H07.4.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P009G03."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Nipponbare;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RA "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 1, PAC
 RT clone: P0030H07."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF002522; BAB3621.1; -.
 DR EMBL; AF003045; BAB44042.1; -.
 DR Gramineae; Q9LGI5; -.
 DR GO; GO:000524; F1ATP binding; IEA.
 DR GO; GO:0004674; F1protein serine/threonine kinase activity; IEA.
 DR GO; GO:0016740; F1transferase activity; IEA.
 DR GO; GO:0006468; F1protein amino acid phosphorylation; IEA.
 DR InterPro; IPR001611; LRR.
 DR InterPro; IPR007091; LRR_plant.
 DR InterPro; IPR00719; Prot_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR Pfam; PF00560; pkinase; 1.
 DR PRINTS; PR00019; LEURICHRPT.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
 DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
 KW Hypothetical protein; ATP-binding; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 1065 AA; 115226 MW; 9C221B26F64AB551 CRC64;

Query Match 0.9%; Score 9; DB 10; Length 1065;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 917 LRELDLSFN 925

```

>
Db      125 LRELDLSFN 133
|||||
RESULT 10
Q86W24 PRELIMINARY; PRT; 1093 AA.
ID Q86W24
AC Q86W24;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE NALP14.
GN NALP14.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22451042; PubMed=12563287;
RA Tschopp J., Martinon F., Burns K.;
RL "NALPs: a novel protein family involved in inflammation.";
DR EMBL; AY154469; AAO18165.1; -.
DR GO; GO:0005743; C:mitochondrial inner membrane; IEA.
DR GO; GO:0005488; F:binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR002194; Chaperonin_TCP-1.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh.sub.
DR InterPro; IPR001993; Mitoch carrier.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR SMART; SM00368; LRR_RI; 12.
DR PROSITE; PS0824; DAPIN; 1.
DR PROSITE; PS00215; MITOCH_CARRIER; 1.
DR PROSITE; PS0837; NACHT; 1.
DR PROSITE; PS00995; TCPI_3; 1.
SQ SEQUENCE 1093 AA; 124732 MW; 124EEACE22A11D6F CRC64;

Query Match 0.9%; Score 9; DB 4; Length 1093;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 711 LRHPNCKLQ 719
Db 697 LRHPNCKLQ 705
|||||
RESULT 11
Q86UB5 PRELIMINARY; PRT; 1375 AA.
ID Q86UB5
AC Q86UB5;
DT 01-JUN-2003 (TREMELrel. 24, Created)
DT 01-JUN-2003 (TREMELrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Similar to death effector filament-forming Ced-4-like apoptosis
DE protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallamy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzyszewski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
RT *Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA Strausberg R.;
RL Submitted (May-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC051787; AAH51787.1; -.
DR GO; GO:0003793; F:defense/immunity protein activity; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007091; LRR_RNinh.
DR InterPro; IPR003590; LRR_RNinh.sub.
DR InterPro; IPR007111; NACHT_NTPase.
DR InterPro; IPR004020; PAAD_DAPIN_dom.
DR Pfam; PF00560; LRR; 2.
DR Pfam; PF02758; PAAD_DAPIN; 1.
DR PRINTS; PR00364; DISEASERESIST.
DR SMART; SM00368; LRR_RI; 5.
DR PROSITE; PS0824; DAPIN; 1.
DR PROSITE; PS0837; NACHT; 1.
SQ SEQUENCE 1375 AA; 154869 MW; 7C834D47BBD490FE CRC64;

Query Match 0.9%; Score 9; DB 4; Length 1375;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 QGAAGIGKS 197
Db 333 QGAAGIGKS 341
|||||
RESULT 12
Q8XA86 PRELIMINARY; PRT; 43 AA.
ID Q8XA86
AC Q8XA86;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-OCT-2002 (TREMELrel. 22, Last annotation update)
DE Hypothetical protein CT2282.
GN CT2282.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=JLS / ATCC 49652 / DSM 12025;
RC MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwin M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Nierman W.C., Feldblyum T.V., Hansen C.L., Brenner M., Shea T.P., Parksey D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
photosynthetic, anaerobic, green-sulfur bacterium.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR EMBL; AE012973; AA73495.1; -.
DR TIGR; CT2282; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 43 AA; 4973 MW; 796C897A0A19RCF0 CRC64;

Query Match 0.8%; Score 8; DB 16; Length 43;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1002 KLRVLMWF 1009
Db 22 KLRVLMWF 29
|||||

RESULT 13
Q8ZB55 PRELIMINARY; PRT; 87 AA.
AC Q8ZB55;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE BoliA-like protein (Hypothetical protein).
GN YPO3570 OR Y0141.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Felwell T., Hamlin N., Holroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL; AJ414157; CAC92799.1; -.
DR EMBL; AE013614; AA883735.1; -.
DR PIR; AC0434; AC0434.
DR InterPro; IPR002634; BOLA.
DR PFam; PF01722; BOLA; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 87 AA; 9962 MW; 1403F68ACF7A254A CRC64;

Query Match 0.8%; Score 8; DB 16; Length 87;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 604 SHFQVIVV 611
Db 29 SHFQVIVV 36
|||||

RESULT 14
Q8FR50 PRELIMINARY; PRT; 93 AA.
ID Q8FR50
AC Q8FR50;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN CE0916.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005217; BAC17726.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 93 AA; 10616 MW; 9B610E35B28AB75B CRC64;

Query Match 0.8%; Score 8; DB 16; Length 93;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 ELKPSFHD 278
Db 5 ELKPSFHD 12
|||||

RESULT 15
Q9ZLP7 PRELIMINARY; PRT; 123 AA.
ID Q9ZLP7
AC Q9ZLP7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative flagellar motor switch protein.
GN FLIN OR JHP0531.
OS Helicobacter pylori J99 (Campylobacter pylori J99).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=85963;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori.";
RL Nature 397:176-180(1999).
DR EMBL; AE001485; AAD06097.1; -.
DR PIR; D71922; D71922.
DR GO; GO:0009425; C:flagellar basal body (sensu Bacteria); IEA.
DR GO; GO:0003774; F:motor activity; IEA.
DR GO; GO:0006935; P:chemotaxis; IEA.
DR GO; GO:0001539; P:ciliary/flagellar motility; IEA.
DR InterPro; IPR001172; Flagellar_FLIN.
DR InterPro; IPR001543; Spoa.
DR PFam; PF01052; Spoa; 1.
DR PRINTS; PR00956; FLGMOTORFLIN.
DR ProDom; PD001777; Spoa; 1.
KW Complete proteome.
SQ SEQUENCE 123 AA; 13966 MW; 6ED16EF028BF1CC1 CRC64;

Query Match 0.8%; Score 8; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 13 LSTYLEEL 20
| | | | |
Db 27 LSTYLEEL 34

Search completed: July 30, 2004, 14:03:46
Job time : 49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 00:15:25 ; Search time 11825 Seconds
(without alignments)
11391.979 Million cell updates/sec

Title: US-10-781-294-23
Perfect score: 3108
Sequence: 1 atgtctacgaacgcaggcag.....attggacattgctgctga 3108

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rtd.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3108	100.0	3108	6	AX417214	AX417214 Sequence
2	3057	98.4	3300	6	AX459869	AX459869 Sequence
3	3017	97.1	3731	9	AY116204	AY116204 Homo sapi
4	3001	96.6	3507	9	BC028069	BC028069 Homo sapi
5	2997.8	96.5	3827	9	AY095146	AY095146 Homo sapi
6	2997.8	96.5	3827	9	AY154467	AY154467 Homo sapi
7	2759	88.8	3563	9	AY116205	AY116205 Homo sapi
8	2644.2	85.1	3466	6	AX833594	AX833594 Sequence
9	2644.2	85.1	3466	9	AK095460	AK095460 Homo sapi
10	2470	79.5	3395	9	AY116206	AY116206 Homo sapi
11	2417	77.8	3221	9	AY116207	AY116207 Homo sapi
12	1703	54.8	147330	9	AC008753	AC008753 Homo sapi
13	1681.8	54.1	2158	6	AX575503	AX575503 Sequence
14	1616	52.0	4931	6	AX684291	AX684291 Sequence
15	1256.2	40.4	2159	10	AY364010	AY364010 Mus muscu
16	1036.2	33.3	223344	2	AC106193	AC106193 Rattus no
17	997.8	32.1	218270	2	AC079499	AC079499 Mus muscu
18	803.6	25.9	4170	9	AF410477	AF410477 Homo sapi
19	800.4	25.8	3102	6	AX299762	AX299762 Sequence
20	800.4	25.8	3579	9	AF468522	AF468522 Homo sapi
21	800.4	25.8	3857	9	AX299760	AX299760 Sequence
22	800.4	25.8	3857	9	AF420469	AF420469 Homo sapi
23	783.4	25.2	3913	10	AF486632	AF486632 Mus muscu
24	783.2	25.2	2934	9	AY422168	AY422168 Homo sapi
25	700.2	22.5	3828	9	AF427617	AF427617 Homo sapi
26	698.6	22.5	3237	9	AY0511283	AY0511283 Homo sapi
27	639	20.6	1958	9	AY0511283	AY0511283 Homo sapi
28	639	20.6	2524	6	AX417249	AX417249 Sequence
29	637.4	20.5	2528	9	AF418985	AF418985 Homo sapi
30	637.4	20.5	212426	9	AC104335	AC104335 Homo sapi
31	626.6	20.2	2546	6	AX780309	AX780309 Sequence
32	625.4	20.1	2546	9	AK027194	AK027194 Homo sapi
33	595.6	19.2	2494	9	AF231021	AF231021 Homo sapi
34	589	19.0	157300	10	AL592522	AL592522 Mouse DNA
35	589	19.0	241714	2	AC083834	AC083834 Mus muscu
36	586.6	18.9	229028	2	AC123316	AC123316 Rattus no
37	586.6	18.9	229793	2	AC098184	AC098184 Rattus no
38	580.2	18.7	223791	2	AC084042	AC084042 Mus muscu
39	475.4	15.3	3282	9	AY154469	AY154469 Homo sapi
40	470.4	15.1	1557	6	AX459897	AX459897 Sequence
41	402.8	13.0	3489	6	AX478549	AX478549 Sequence
42	402.8	13.0	3830	6	AX704823	AX704823 Sequence
43	402.8	13.0	3885	9	AY154460	AY154460 Homo sapi
44	402.8	13.0	3926	6	AX704821	AX704821 Sequence
45	402.8	13.0	5859	6	AX459881	AX459881 Sequence

ALIGNMENTS

RESULT 1
AX417214
LOCUS AX417214 3108 bp DNA linear PAT 14-JUN-2002
DEFINITION Sequence 23 from Patent WO0226780.
ACCESSION AX417214
VERSION AX417214.1 GI:21449784
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Read, J.C., Godzik, A., Chu, Z.L., Pawlowski, K., Fiorentino, L.,
Ariza, M.E. and Stehlik, C.
TITLE Paad domain-containing polypeptides, encoding nucleic acids, and

Pred. No. is the number of results predicted by chance to have a

methods of use		
JOURNAL	Patent: WO 0226780-A 23 04-APR-2002;	
FEATURES	The Burnham Institute (US)	
source	Location/Qualifiers	
	1. .3108	
	/organism="Homo sapiens"	
	/mol_type="unassigned DNA"	
	/db_xref="taxon:9606"	
CDS	1..3108	
	/note="unnamed protein product"	
	/codon_start=1	
	/protein_id="CAD35284.1"	
	/db_xref="GI:21449785"	
	/db_xref="REMBL:CAD35284"	
	/translation="MLTAGRDGLCRISTYLELEAVELEKFKLYLGATATGELGKPI	
	WGSMEKAGPLEMAQLLIITHTGPBEARLALSTPERINRDLWEGGREDLVRPOEITY	
	RDYVRRFRMEDNARLBCVNSHRYTRLLLVKHSNPMQVQQLDITGRGHART	
	GHAQSPKIEITLPEPPERPEPRVTVMGAAGIGKSMIAKVMWDADKGLFGFRD	
	YLFYINCRNQSAETCSMDLIPSCWPEPSAPLQELIRPERLLFIIDFDELKPSF	
	HDPGWCLQWEAKRPTLELLSLIRKLLPELSLLITRPTALEKHLLEPRHVE	
	ILGFSEARKEYFYKYPHNAEQAGOVNVRDNEPLFTMCFVPLVWVCTCLOQLE	
	GGILLRQTSRTTAVVMYLLSLMOPKPGAPRLQPPNORGLCSLAADGLWNOKLFE	
	EQDLRKGLDGEDVSALNNNI FQDINCERYYSFIHLSFQEPFAAMYIILDEGEBA	
	GPQDVVRLUTEYAFSPSEALTSRELFGLNNEETSHLEKSLCWKVSPIKMDLLQ	
	MIQSKAGDSGLTQGGSEFFSCLYEIOBEFFIQOALSHFQVIVVNSIAKMEHWSS	
	FCLKRCRSQVLYLGYATYADGEDRARCAGATLLVQLRPERTVILSDAYSEHLAA	
	LCINPNLIELSLVNALSGRVKLLCOGLRHPNCKIQLRLKRCRISSACEIDLAA	
	IANKNLTRMDLSGNGVPGVMWLLCEGLRHPQCRLOMQLRKCOLESACQEMASVLG	
	TNPHVELDITGNALDGLRLLCQGLRHPVCRRLTLWLKRLTAACBELASTLSV	
	NQSLRELDLSNEIGDILGVLILLCEGLRHPTKLGQTLRIGICRIGSAAECGLSVVQAN	
	HNRLDELSPNDLMDGLWLEAGLQHPACRLQKMLDSCGLTAKACENLYFTTGINO	
	LDLGLYLTNNALDGTGVRLLCKRLSHPGCKLRVLWLPQMDLNKMTSHRLAALRVTKY	
	LDIGC"	
ORIGIN		
Query Match	100.0%; Score 3108; DB 6; Length 3108;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 3108; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGCTAGAACCGCAGGAGGAGGCTCTGTCGCTGTCACCTACTCTGGAAGACTC	60
DB	1 ATGCTAGAACCGCAGGAGGAGGCTCTGTCGCTGTCACCTACTCTGGAAGACTC	60
QY	61 GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGACCGGACAGAGACTGGAGAA	120
DB	61 GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGACCGGACAGAGACTGGAGAA	120
QY	121 GGGAAGATCCCTGGGAGGAGCATGAGAGAGCCGGTCCCTGGAAATGGCCAGCTGCTC	180
DB	121 GGGAAGATCCCTGGGAGGAGCATGAGAGAGCCGGTCCCTGGAAATGGCCAGCTGCTC	180
QY	181 ATCAACCACTTGGGACAGAGAGGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA	240
DB	181 ATCAACCACTTGGGACAGAGAGGCTGGAGGTTGGCTCTCAGCACCTTTGAGCGGATA	240
QY	241 AACAGGAGGACCTGTGGGAGAGAGGACAGAGAGGAGGACCTGGTGGAGGATCCCGAGAA	300
DB	241 AACAGGAGGACCTGTGGGAGAGAGGACAGAGAGGAGGACCTGGTGGAGGATCCCGAGAA	300
QY	301 ACCTACAGGAGCTATGTCCGAGAGAAATTCGGGCTCATGGAAGACCGCAATGGCGGCTTA	360
DB	301 ACCTACAGGAGCTATGTCCGAGAGAAATTCGGGCTCATGGAAGACCGCAATGGCGGCTTA	360
QY	361 GGGGAATGTGCAACCTCAGCACCGGTACACCGGCTCCCTGCTGGTGAAGAGGACACTCA	420
DB	361 GGGGAATGTGCAACCTCAGCACCGGTACACCGGCTCCCTGCTGGTGAAGAGGACACTCA	420
QY	421 AACCCCATGCAAGTCCACAGCAGCTTCTGGACACAGGCGGGGACACGCGAGGACCGTG	480
DB	421 AACCCCATGCAAGTCCACAGCAGCTTCTGGACACAGGCGGGGACACGCGAGGACCGTG	480
QY	481 GGACACCAAGGCTAGGCCCCATCAGATAGAGACCTCTTTTGGCCAGAGGAGGCGCCC	540
DB	481 GGACACCAAGGCTAGGCCCCATCAGATAGAGACCTCTTTTGGCCAGAGGAGGCGCCC	540
QY	541 GAGCCACCGCGACCGTGGTTCATGCAAGGCGCGCAGGATAGGCAAGTCTCATCTGGCA	600
DB	541 GAGCCACCGCGACCGTGGTTCATGCAAGGCGCGCAGGATAGGCAAGTCTCATCTGGCA	600
QY	601 CACAAGGTGATGTGGACTGGGCGGACGGAAGCTTTCACAGGACAGATTGATATCTC	660
DB	601 CACAAGGTGATGTGGACTGGGCGGACGGAAGCTTTCACAGGACAGATTGATATCTC	660
QY	661 TTCTACATCAACTGCAGGAGATGAACAGAGTGCACGGAATGACATGACATGACAGACTC	720
DB	661 TTCTACATCAACTGCAGGAGATGAACAGAGTGCACGGAATGACATGACATGACAGACTC	720
QY	721 ATCTTCAGCTGCTGGCCTGAGCCAGCGCTCTCCAGAGACTCATCCGAGTTCCCGAG	780
DB	721 ATCTTCAGCTGCTGGCCTGAGCCAGCGCTCTCCAGAGACTCATCCGAGTTCCCGAG	780
QY	781 CGCCTCCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCCACGATCTCAG	840
DB	781 CGCCTCCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCCACGATCTCAG	840
QY	841 GGACCTCTGCTGCTCTGGGAGGAGAAACGCGCCACGAGCTCTTCTTAAACAGCTTA	900
DB	841 GGACCTCTGCTGCTCTGGGAGGAGAAACGCGCCACGAGCTCTTCTTAAACAGCTTA	900
QY	901 ATTCGGAAGAGCTGCTCTCTGCTGAGCTATCTTTGCTCATCACACGCGCCACGCTTTG	960
DB	901 ATTCGGAAGAGCTGCTCTCTGCTGAGCTATCTTTGCTCATCACACGCGCCACGCTTTG	960
QY	961 GAGAAGTTCACCTGCTGCTGGAGCACCCAGGAGCTGTGGAGATCTGGGCTTCTCTGAG	1020
DB	961 GAGAAGTTCACCTGCTGCTGGAGCACCCAGGAGCTGTGGAGATCTGGGCTTCTCTGAG	1020
QY	1021 GCAGAAAGGAGGATCTTCTACAGTATTTTCCAAATGCAGAGCGGCGGCGCAAGTC	1080
DB	1021 GCAGAAAGGAGGATCTTCTACAGTATTTTCCAAATGCAGAGCGGCGGCGCAAGTC	1080
QY	1081 TTCAATTAACGTGAGGAGCAACAGGCTCTCTTCCACCATGTGCTTCTGCTGCTGCTGTC	1140
DB	1081 TTCAATTAACGTGAGGAGCAACAGGCTCTCTTCCACCATGTGCTTCTGCTGCTGCTGTC	1140
QY	1141 TGGTGTGTGTACTGCTCAGGAGCTGAGAGGTTGGGGGCTGTTGAGACAGAGC	1200
DB	1141 TGGTGTGTGTACTGCTCAGGAGCTGAGAGGTTGGGGGCTGTTGAGACAGAGC	1200
QY	1201 TCCAGGACCACTGCTGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
DB	1201 TCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1260
QY	1261 GGGGCCCCCGCTCCAGCCCCCAGGAGGAGGTTGTGCTCTTGGCGGAGAT	1320
DB	1261 GGGGCCCCCGCTCCAGCCCCCAGGAGGAGGTTGTGCTCTTGGCGGAGAT	1320
QY	1321 GGGCTCTGGAATCAGAAATCCTATTTGAGAGGAGGAGCTCCGGAAGACAGGCTAGAC	1380
DB	1321 GGGCTCTGGAATCAGAAATCCTATTTGAGAGGAGGAGCTCCGGAAGACAGGCTAGAC	1380
QY	1381 GGGGAAGAGCTCTGCTGCTTCTCAACATGAACATCTTCCAGAGGAGATCAACTGTGAG	1440
DB	1381 GGGGAAGAGCTCTGCTGCTTCTCAACATGAACATCTTCCAGAGGAGATCAACTGTGAG	1440
QY	1441 AGGTACTACAGCTTCACTTCACTTGGAGTTCCAGGAATCTTTTGGAGCTATGTACTATATC	1500
DB	1441 AGGTACTACAGCTTCACTTCACTTGGAGTTCCAGGAATCTTTTGGAGCTATGTACTATATC	1500
QY	1501 CTGAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1560
DB	1501 CTGAGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1560
QY	1561 TACGCGTTTCTGAAAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620
DB	1561 TACGCGTTTCTGAAAGGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1620

QY	1621	AACGAGGACACGAGGACCACTCGAGAGAAGTCTCTCTGCTGGAAGTCTCTCGCCGACATC	1680
DB	1621	AACGAGGACACGAGGACCACTCGAGAGAAGTCTCTCTGCTGGAAGTCTCTCGCCGACATC	1680
QY	1681	AAGATGGACCTGTTGCGAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGTCCACCCCTCGAG	1740
DB	1681	AAGATGGACCTGTTGCGAGTGGATCCAAAGCAAAAGCTCAGAGCGACGGTCCACCCCTCGAG	1740
QY	1741	CAGGCGCTCCTTGGAGTCTTCTCAGCTGCTGTACAGATCCAGGAGGAGGAGTTTATCCAG	1800
DB	1741	CAGGCGCTCCTTGGAGTCTTCTCAGCTGCTGTACAGATCCAGGAGGAGGAGTTTATCCAG	1800
QY	1801	CAGGCCCTGAGCCATCTCCAGTGATCTGCTGTCAGCAACATTCGCTCCAGATGGAGCAC	1860
DB	1801	CAGGCCCTGAGCCATCTCCAGTGATCTGCTGTCAGCAACATTCGCTCCAGATGGAGCAC	1860
QY	1861	ATGGTCTCTCTGTTCTGTCTGAAGCGTGCAGAGCGGCCAGGTGCTGTCATCTGTATGGC	1920
DB	1861	ATGGTCTCTCTGTTCTGTCTGAAGCGTGCAGAGCGGCCAGGTGCTGTCATCTGTATGGC	1920
QY	1921	GCCACCTACAGCGCGGACGSGGAGACCGCGGAGGTGCTCCGAGAGCGCACACGCTG	1980
DB	1921	GCCACCTACAGCGCGGACGSGGAGACCGCGGAGGTGCTCCGAGAGCGCACACGCTG	1980
QY	1981	TTGGTGCAGCTCAGACCAGAGGACCGTCTCTGCTGGACGCTACAGTGAACATCTGGCA	2040
DB	1981	TTGGTGCAGCTCAGACCAGAGGACCGTCTCTGCTGGACGCTACAGTGAACATCTGGCA	2040
QY	2041	GCGGCCCTGTGCACCAATCCAACTGATAGAGTGTCTGTGACGCTTACAGTGAACATCTGGCA	2100
DB	2041	GCGGCCCTGTGCACCAATCCAACTGATAGAGTGTCTGTGACGCTTACAGTGAACATCTGGCA	2100
QY	2101	AGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC	2160
DB	2101	AGCCGGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGCAAACTTCAGAAC	2160
QY	2161	CTGAGGCTGAAGAGGTCCGCATCTCCAGACTCAGCCTCGAGGACCTCTCTCGAGCTCTC	2220
DB	2161	CTGAGGCTGAAGAGGTCCGCATCTCCAGACTCAGCCTCGAGGACCTCTCTCGAGCTCTC	2220
QY	2221	ATAGCCCAATAGAAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATCCAGGC	2280
DB	2221	ATAGCCCAATAGAAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATCCAGGC	2280
QY	2281	ATGATGCTGCTTTTCGAGGGCCTCGCGCATCCCGAGTCCCGAGGCTGAGATGATTCAGTTG	2340
DB	2281	ATGATGCTGCTTTTCGAGGGCCTCGCGCATCCCGAGTCCCGAGGCTGAGATGATTCAGTTG	2340
QY	2341	AGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTGCTCGGACCAAC	2400
DB	2341	AGGAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGTGCTCGGACCAAC	2400
QY	2401	CCACATCTGGTTGAGTTGGACTTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTA	2460
DB	2401	CCACATCTGGTTGAGTTGGACTTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTA	2460
QY	2461	CTATGCCAGGCACTGAGGACCCAGTCTGACACTACGACTTGTGGCTGAGATCTGCG	2520
DB	2461	CTATGCCAGGCACTGAGGACCCAGTCTGACACTACGACTTGTGGCTGAGATCTGCG	2520
QY	2521	CGCCTCACTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTG	2580
DB	2521	CGCCTCACTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTG	2580
QY	2581	AGAGAGCTGACCTGAGCTGAATGAGTGGGGAACCTCGGGGTGCTGCTGTGTGAG	2640
DB	2581	AGAGAGCTGACCTGAGCTGAATGAGTGGGGAACCTCGGGGTGCTGCTGTGTGAG	2640
QY	2641	GGCCTCAGGCATCCACGCTGACGCTCAGACCCCTCGGGTTCGGGATCTGCCGCTGGGC	2700
DB	2641	GGCCTCAGGCATCCACGCTGACGCTCAGACCCCTCGGGTTCGGGATCTGCCGCTGGGC	2700

QY	2701	TCGCGCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGGAGCTG	2760
Db	2701	TCTGCGCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAACCTCCGGGAGCTG	2760
QY	2761	GACTTGAGTTTTCAACGACTCTGGAGACTGGGGCCTGTGGTTGCTGGCTGAGGGGCTGCAA	2820
Db	2761	GACTTGAGTTTTCAACGACTCTGGAGACTGGGGCCTGTGGTTGCTGGCTGAGGGGCTGCAA	2820
QY	2821	CATCCCGGCTGCAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCTCCAGACCACAGCCAAAGGCT	2880
Db	2821	CATCCCGGCTGCAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCTCCAGACCACAGCCAAAGGCT	2880
QY	2881	TGTGAAGATCTTTACTTCACCTGGGATCAACAGACCTTTGACCGACCTTTTACTGACC	2940
Db	2881	TGTGAAGATCTTTACTTCACCTGGGATCAACAGACCTTTGACCGACCTTTTACTGACC	2940
QY	2941	AACAACGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGCAAGCGCTGAGCCATCTCTGGC	3000
Db	2941	AACAACGCCCTTAGGGGACAGAGTGTCCGACTGCTTTGCAAGCGCTGAGCCATCTCTGGC	3000
QY	3001	TGAAAACCTCCGAGTCCCTCTGTTATTTTGGGATGACCTGAATAAATGACCCACAGTAGG	3060
Db	3001	TGAAAACCTCCGAGTCCCTCTGTTATTTTGGGATGACCTGAATAAATGACCCACAGTAGG	3060
QY	3061	TGGCAGCGCTTCGAGTAACAAAACCTTATTGGACATTGGCTGCTGA	3108
Db	3061	TGGCAGCGCTTCGAGTAACAAAACCTTATTGGACATTGGCTGCTGA	3108
<hr/>			
RESULT 2			
AX459869			
LOCUS			
AX459869			
DEFINITION			
Sequence 12 from Patent WO0240668.			
ACCESSION			
AX459869			
VERSION			
AX459869.1 GI:21725643			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Homo sapiens			
REFERENCE			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS			
1 Teschopp,J. and Martinon,F.			
TITLE			
Proteins and dna sequences underlying these proteins used for treating inflammations			
JOURNAL			
Patent: WO 0240668-A 12 23-MAY-2002;			
Apotech Research and Development Ltd. (CH)			
FEATURES			
Location/Qualifiers			
source			
1..3300			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
/note="NALP3/PY5.dna"			
<hr/>			
ORIGIN			
<hr/>			
Query Match			
Best Local Similarity 98.4%; Score 3057; DB 6; Length 3300;			
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
<hr/>			
QY	1	ATGCTACGAAACCGCAGCGAGGACGGCCTCTGCGCTGTCCACCTACTTGGAGAATC	60
Db	1	ATGCTACGAAACCGCAGCGAGGACGGCCTCTGCGCTGTCCACCTACTTGGAGAATC	60
QY	61	GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGGACCGCACAGAGCTGGAGAA	120
Db	61	GAGGCTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGGACCGCACAGAGCTGGAGAA	120
QY	121	GGCAAGATCCCTGGGGAAGCATGAGAGGCGCGTCCCTGGAAATGGCCAGCTGCTC	180
Db	121	GGCAAGATCCCTGGGGAAGCATGAGAGGCGCGTCCCTGGAAATGGCCAGCTGCTC	180
QY	181	ATCACCACCTTGGGCCACAGAGAGGCTTGGCTTCTAGCACTTTGACGCGATA	240
Db	181	ATCACCACCTTGGGCCACAGAGAGGCTTGGCTTCTAGCACTTTGACGCGATA	240

QY	241	AA	CAGAAAGGACCTGTGGGAGAGACAGACAGAGAGGACCTGGTGAAGGATCCCCAGGAA	300
Db	241	AA	CAGAAAGGACCTGTGGGAGAGAGACAGAGAGGACCTGGTGAAGGATCCCCAGGAA	300
QY	301	AC	TACAGGACATATGTCGACAGAAATTCGGCTCATGGAAGACCGCAATCGCGCTTA	360
Db	301	AC	TACAGGACATATGTCGACAGAAATTCGGCTCATGGAAGACCGCAATCGCGCTTA	360
QY	361	GGG	GAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCTGTGTGTAAGGAGCACTCA	420
Db	361	GGG	GAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCTGTGTGTAAGGAGCACTCA	420
QY	421	AA	CCCCATGACGCTCAGACGACGCTTCTGGAACAGCCGGGACACCGAGGACCGTG	480
Db	421	AA	CCCCATGACGCTCAGACGACGCTTCTGGAACAGCCGGGACACCGAGGACCGTG	480
QY	481	GG	ACACAGGCTAGCCCCATCAAGATAGACACCTCTTTGAGCAGACAGAGGACGCCCC	540
Db	481	GG	ACACAGGCTAGCCCCATCAAGATAGACACCTCTTTGAGCAGACAGAGGACGCCCC	540
QY	541	GAG	CCACCGGCACCGTGGTATGCAAGGCGCGCAGGATAGCGAAGTCCATGCTGGCA	600
Db	541	GAG	CCACCGGCACCGTGGTATGCAAGGCGCGCAGGATAGCGAAGTCCATGCTGGCA	600
QY	601	CAC	AAGTGATGCTGGACTGGGCGGACGGGAGCTCTTCCAGGACAGATTTGATTATCTC	660
Db	601	CAC	AAGTGATGCTGGACTGGGCGGACGGGAGCTCTTCCAGGACAGATTTGATTATCTC	660
QY	661	TT	TACATCAACTGCAGGGAGATGAACAGAGTGCACGGAATGCAGATGCAGACCTC	720
Db	661	TT	TACATCAACTGCAGGGAGATGAACAGAGTGCACGGAATGCAGATGCAGACCTC	720
QY	721	AT	CTTCAGCTGTGSCCTGAGCCACGCGCCTCTCCAGGAGCTCATCCGAGTTCGCGAG	780
Db	721	AT	CTTCAGCTGTGSCCTGAGCCACGCGCCTCTCCAGGAGCTCATCCGAGTTCGCGAG	780
QY	781	CG	CTCTTTTCATCATGACGGCTTCGATGAGCTCAAGCTTCTTCCAGCATCTCAG	840
Db	781	CG	CTCTCTTTTCATCATGACGGCTTCGATGAGCTCAAGCTTCTTCCAGCATCTCAG	840
QY	841	GG	ACCTGGTCTCTGCTGGAGAGAAACGGCCACGAGCTGCTTTAAACAGCTTA	900
Db	841	GG	ACCTGGTCTCTGCTGGAGAGAAACGGCCACGAGCTGCTTTAAACAGCTTA	900
QY	901	ATT	CGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACGCGCCACCGCTTG	960
Db	901	ATT	CGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACGCGCCACCGCTTG	960
QY	961	GAG	AGCTCCACCGTCTGTGAGCACCGCAGGATGTGGATCTCTGGCTTCTCTGAG	1020
Db	961	GAG	AGCTCCACCGTCTGTGAGCACCGCAGGATGTGGATCTCTGGCTTCTCTGAG	1020
QY	1021	GC	AGAGGAAGAAATACCTTCTACAAGTATTTCCACAATCAGAGCAGCGCGCAAGTC	1080
Db	1021	GC	AGAGGAAGAAATACCTTCTACAAGTATTTCCACAATCAGAGCAGCGCGCAAGTC	1080
QY	1081	TT	CAATTAAGTGTAGGAGCAACGAGGCTCTTTCAACATGTGCTTCGTCGCCCTGGTGTGC	1140
Db	1081	TT	CAATTAAGTGTAGGAGCAACGAGGCTCTTTCAACATGTGCTTCGTCGCCCTGGTGTGC	1140
QY	1141	TG	GGTGTGTATCTGCTCAGCAGCTGAGGAGTGGGGCTCTTGAGACAGAGC	1200
Db	1141	TG	GGTGTGTGTATCTGCTCAGCAGCTGAGGAGTGGGGCTCTTGAGACAGAGC	1200
QY	1201	TC	CAGACACCACTGATGTATGCTCTACCTGCTGAGTCTGATGCAACCCCAAGCG	1260
Db	1201	TC	CAGACCAACCACTGATGTATGCTCTACCTGCTGAGTCTGATGCAACCCCAAGCG	1260
QY	1261	GG	GGCCCGCGCTCCAGCCCCCAACACAGAGAGGTTGTGCTCTTGGCGCAGAT	1320
Db	1261	GG	GGCCCGCGCTCCAGCCCCCAACACAGAGAGGTTGTGCTCTTGGCGCAGAT	1320

CDS

```
177..3362
/map="19q13.4"
/note="PYPAF7; PYRIN-containing nucleotide-binding site/
leucine-rich repeat family member"
/codon_start=1
/product="PYRIN-containing APAF1-like protein 7"
/protein_id="AAM18227.1"
/db_xref="GI:21314907"
```

RLQPPNQRGLCSLAADGLMKNKILFEEDLRKHGLDGVSAFLNMNIPQKDINCER
YYGFIHLSFOEPFAAMVYIIDEDEGGAGDPDVRILLTEYAFSEFSLALTSRFLGL
LNETSRHSLCWKVSPHIKMDLLQWIKQAQSDGSLTQOGLSEFFSLCYBIOEBE
FIQOALSHFOVIVVSNIAKMEHWSFCLKRCRQAQVHLXGATVYSDGEDRARC
GATHLVQDLPRTVLDAYSEDLAAALCTNPNIELSLYRNALGSRGVKLLCOGLRHP
NCKLQNLRLKRCRISSEADLSAALIANKRMDSLGNVDFPGMMLLCEBGLRHP
CLQMTQLRKCQLESACQEMASVLTGNPHLVELDTGNALDELGLRLCQGLRHPV
RLRLTLGICRLTAAACDELASTLSVNSQSLRELDLSNELGDLGLVLLCEGLRHPCK
LQTLRLGICRLTAAACDELASTLSVNSQSLRELDLSFNDLGDWGLWLLAEGLRHPCKL
OKLWLDSCGLTAKACENLYETLGINOTLDTLXLTNNALGDTGVRLLLCKRLSHPGCKLR
VLWLFMDLNMTHSRLLAALRVKPYLDIGC"

ORIGIN

Query Match	96.5%;	Score	2997.8;	DB	9;	Length	3827;
Best Local Similarity	97.3%;	Pred. No.	0;				
Matches	3103;	Conservative	0;	Mismatches	2;	Indels	84;
Gaps	2;						
QY	1	ATGCTACGAACCGCAGCAGGACGGCTCTGTGCGCTGTCCACCTTACTTGGAAAGACTC	60				
Db	177	ATGCTACGAACCGCAGCAGGACGGCTCTGTGCGCTGTCCACCTTACTTGGAAAGACTC	236				
QY	61	GAGGCTGTGAACTGAAGAAGTTCAAGTTATATCTCGGGACCGCGACAGAGCTGGAGAA	120				
Db	237	GAGGCTGTGAACTGAAGAAGTTCAAGTTATATCTCGGGACCGCGACAGAGCTGGAGAA	296				
QY	121	GGCAAGATCCCTCGGGGAAGCATGGAGAGCCGGTCCCTGGAATGCCAGCTGCTC	180				
Db	297	GGCAAGATCCCTCGGGGAAGCATGGAGAGCCGGTCCCTGGAATGCCAGCTGCTC	356				
QY	181	ATCACCACTTCGGGCGCAGAGAGGCTTGAAGTTGGCTTCAGACCTTTGAGCGGATA	240				
Db	357	ATCACCACTTCGGGCGCAGAGAGGCTTGAAGTTGGCTTCAGACCTTTGAGCGGATA	416				
QY	241	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGG-----	288				
Db	417	AACAGGAAGACCTGTGGGAGAGGACAGAGAGGACCTGTGTGAGGATACCCACCT	476				
QY	289	-----	288				
Db	477	GGTGGCGGCTCTCACTTGGGAACCAAGTCAACATGCTTCTGTGAAGTCTCTTTGTCACT	536				
QY	289	-----GATCCCGAGGAACCTACAGGACCTATGTCCGCGAGGAATTCGGCTCATG	339				
Db	537	CCAAAGAAAGATCCCGAGGAACCTACAGGACCTATGTCCGCGAGGAATTCGGCTCATG	596				
QY	340	GAAGACCGCAATCGCGCTTAGGGGAATGTCAAACCTCAGCCACCGGTACACCGGCTC	399				
Db	597	GAAGACCGCAATCGCGCTTAGGGGAATGTCAAACCTCAGCCACCGGTACACCGGCTC	656				
QY	400	CTGCTGGTGAAGAGCACTCAAAACCCCATGAGGTCCAGCAGCAGCTTCTGGACACAGGC	459				
Db	657	CTGCTGGTGAAGAGCACTCAAAACCCCATGAGGTCCAGCAGCAGCTTCTGGACACAGGC	716				
QY	460	CGGGGACACCGGAGGACCGTGGGACACCAAGGCTAGCCCATCAAGATAGAGACCTCTTT	519				
Db	717	CGGGGACACCGGAGGACCGTGGGACACCAAGGCTAGCCCATCAAGATAGAGACCTCTTT	776				
QY	520	GAGCCAGACGAGGAGCGCCCGAGCCACCGCGACCTGGTTCATATGAAGCGCGGACAGG	579				
Db	777	GAGCCAGACGAGGAGCGCCCGAGCCACCGCGACCTGGTTCATATGAAGCGCGGACAGG	836				
QY	580	ATAGGCAAGTCCATGCTGGGACACAAAGGTGTATGCTGGAGTGGGCGGACGGAGCTTTC	639				
Db	837	ATAGGCAAGTCCATGCTGGGACACAAAGGTGTATGCTGGAGTGGGCGGACGGAGCTTTC	896				
QY	640	CAAGGCAAGTTTGATTATCTTCTTCACTCAACTGACGGGAGATGAACCAAGTGCACG	699				
Db	897	CAAGGCAAGTTTGATTATCTTCTTCACTCAACTGACGGGAGATGAACCAAGTGCACG	956				
QY	700	GAATGCAGCATGCAAGACCTCATCTTCACTGCTGGCTGAGCCCGGCGCTCTCCAG	759				
Db	957	GAATGCAGCATGCAAGACCTCATCTTCACTGCTGGCTGAGCCCGGCGCTCTCCAG	1016				

Qy	760	GAGTCTATCCGAGTTCCGAGCGCCTCTTTTATCATCATCAAGCGCTTCGATGAGCTCAAG	819
Db	1017	GAGTCTATCCGAGTTCCGAGCGCCTCTTTTATCATCATCAAGCGCTTCGATGAGCTCAAG	1076
Qy	820	CCCTCTTTCCACCATCTCAGGAGACCTGTGCTCTGTCTGGGAGGAGAAACGGCCCAACG	879
Db	1077	CCCTCTTTCCACCATCTCAGGAGACCTGTGCTCTGTCTGGGAGGAGAAACGGCCCAACG	1136
Qy	880	GAGTCTCTTTTAAACAGCTTAAATTCGGAAGAAGCTGTCTCTGAGCTATCTTTGCTCATC	939
Db	1137	GAGTCTCTTTTAAACAGCTTAAATTCGGAAGAAGCTGTCTCTGAGCTATCTTTGCTCATC	1196
Qy	940	ACCACAGCGCCACCGGCTTTGGAGAGACTCCACGCTGTCTGGAGCAACCCAGCATGTG	999
Db	1197	ACCACAGCGCCACCGGCTTTGGAGAGACTCCACGCTGTCTGGAGCAACCCAGCATGTG	1256
Qy	1000	GAGATCTCTGGGCTTCTCTGAGGAGAGAAAGAAAGTACTTTCTACAAAGTATTTCCACAAT	1059
Db	1257	GAGATCTCTGGGCTTCTCTGAGGAGAGAAAGAAAGTACTTTCTACAAAGTATTTCCACAAT	1316
Qy	1060	GCAGAGCAGCGCGGCCCAAGTCTTCAATTACGTGAGGAGCAACGAGCCCTCTCTTCCACATG	1119
Db	1317	GCAGAGCAGCGCGGCCCAAGTCTTCAATTACGTGAGGAGCAACGAGCCCTCTCTTCCACATG	1376
Qy	1120	TGCTTCGTCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	1179
Db	1377	TGCTTCGTCCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	1436
Qy	1180	GGGGGGCTGTGACAGACAGCTCCAGGACACCACTGCACTGAGTGTACATGCTCTACTGCTG	1239
Db	1437	GGGGGGCTGTGACAGACAGCTCCAGGACACCACTGCACTGAGTGTACATGCTCTACTGCTG	1496
Qy	1240	AGTCTGATGCAACCAAGCGGGGGCCCGCGCTCCAGGCCCCACCCCAACAGAGAGGG	1299
Db	1497	AGTCTGATGCAACCAAGCGGGGGCCCGCGCTCCAGGCCCCACCCCAACAGAGAGGG	1556
Qy	1300	TTGTGCTCTTGGCGGAGATGGCTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG	1359
Db	1557	TTGTGCTCTTGGCGGAGATGGCTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTG	1616
Qy	1360	CTCCGGAAGCAGCGCTAGACGGGGAAGACCTCTCTGCTCTCTCTCAACATGAAATCTTTC	1419
Db	1617	CTCCGGAAGCAGCGCTAGACGGGGAAGACCTCTCTGCTCTCTCTCAACATGAAATCTTTC	1676
Qy	1420	CAGAAGACATCAACTGTGAGAGGTACTACAGCTTCTACCTGAGTCTGAGTCTGAGTCTGAGTCTG	1479
Db	1677	CAGAAGACATCAACTGTGAGAGGTACTACAGCTTCTACCTGAGTCTGAGTCTGAGTCTGAGTCTG	1736
Qy	1480	TTTGCAGCTATGTACTATATCTCTGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1539
Db	1737	TTTGCAGCTATGTACTATATCTCTGAGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	1796
Qy	1540	GTGACAGGCTGTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCTCTGGCACTCAACAGC	1599
Db	1797	GTGACAGGCTGTGTGACCGAGTACCGCTTTTCTGAAAGGAGCTTCTCTGGCACTCAACAGC	1856
Qy	1600	CGCTTCTGTGTGAGACTCCTGAAACAGGAGACCCAGGAGCCACTGGAGAGAGTCTCTGCG	1659
Db	1857	CGCTTCTGTGTGAGACTCCTGAAACAGGAGACCCAGGAGCCACTGGAGAGAGTCTCTGCG	1916
Qy	1660	TGGAGGTCTCGCGCACATCAAGATGGACCTGTGAGTGTGATCCAAAGCAAGCTCAG	1719
Db	1917	TGGAGGTCTCGCGCACATCAAGATGGACCTGTGAGTGTGATCCAAAGCAAGCTCAG	1976
Qy	1720	AGCGACGCTCCACCTGCGAGCGGCTCTTGGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTT	1779
Db	1977	AGCGACGCTCCACCTGCGAGCGGCTCTTGGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCTT	2036
Qy	1780	CAGGAGGAGGAGTTTATCCAGAGGCCCTGAGCCACTTCCAGAGTGTATGCTGTGTGAGCAAC	1839
Db	2037	CAGGAGGAGGAGTTTATCCAGAGGCCCTGAGCCACTTCCAGAGTGTATGCTGTGTGAGCAAC	2096
Qy	1840	ATTGCTCTCAAGATGAGACATGCTCTCTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTGCTGTCTG	1899

[illegible]

Db 3174 TTGACCGACCTTTTACCTGACCAACAAACGCCCTAGGGGACACAGGTGTCCGACTGCTTTTGC 3233

QY 2980 AAGCGGCTGAGCCATCTCTGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACACTG 3039

Db 3234 AAGCGGCTGAGCCATCTCTGGCTGCAAACTCCGAGTCTCTGTTATTTGGATGACACTG 3293

QY 3040 AATAAATGACCCACAGTAGTGTGGCAGCGCTTCGAGTAACAAACCTTATTTTGGACATT 3099

Db 3294 AATAAATGACCCACAGTAGTGTGGCAGCGCTTCGAGTAACAAACCTTATTTTGGACATT 3353

QY 3100 GCCTGCTGA 3108

Db 3354 GCCTGCTGA 3362

RESULT 6

AY154467 3827 bp mRNA linear PRI 20-FEB-2003

LOCUS Homo sapiens NALP12 (NALP12) mRNA, complete cds.

DEFINITION

ACCESSION AY154467

VERSION AY154467.1 GI:28436377

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

REFERENCE

AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

TITLE 1 (bases 1 to 3827)

JOURNAL Tschopp,J., Martinon,F. and Burns,K.

MEDLINE NALPs: a novel protein family involved in inflammation

PUBMED Nat. Rev. Mol. Cell Biol. 4 (2), 95-104 (2003)

22451042

12563287

REFERENCE

AUTHORS 2 (bases 1 to 3827)

TITLE Martinon,F., Hofmann,K. and Tschopp,J.

JOURNAL Direct Submission

FEATURES

source Submitted (25-SEP-2002) Institute of Biochemistry, University of Lausanne, ch. des Boveresses 155, Epalinges, VD 1066, Switzerland

Location/Qualifiers

1..3827

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

1..3827

/gene="NALP12"

177..3362

/gene="NALP12"

/notes="member of the NALP protein family involved in inflammation; contains NACHT, LRR and PYD domains"

/codon_start=1

/product="NALP12"

/protein_id="AA018163.1"

/db_xref="GI:28436378"

/translation="MLRTAGRDGLRCLSTYLEEAVELKKFLYLGTATLGEKGP WSGMEKQGLEMAQLLIHTFGPEEARLALSTFERINKDLWEGOREDLDVDPGG PPSLGNQKLELVSLTPRKDPQETVDYVRKRLMEDRNARLGEVNLSHRYTRL LAYKHSNPMQVQQLDTGRGHATVGHQASPIKIEITLFEDERPSPPTVMQGA APLQELKMLHLLMDWADGKLPQGRFDYLIINCREMNSQATECSMDLIFSCWEP S ELISLITTPTALEKHLRLLHPHVEILGFSEAEKRYFYKYFNAAQAGOVENVR DNEPLFTMCVPLVCWVTCLQOOLEGGILLROT SRTTTA VYMLYLSLMQPKGAP RLQPPNORGLCSLAADGLWQKILFEQDLRKHGLDGEDVSAFNNMNI FQKDINCR YISFHLSPQEPFAMYYILDEGEGAGPDQDVTKLLTYAFSEKSFIALRSLFGLG LNEETSHLEKSCWKVSPHIMDLQWIKQAQSDGSTLQOGLSEFSLCYIOBE E F1QQAHSFQVTVSNIASKMEHWSSFLKRCRSAQVLYHLYGATYSADGIDRAC S GAHTLLVQPERTVLLDAYSEHLAAALCTNPNI,ELSLYRNALGSRGVKLLCGLRHP NCKLOMLRLKRCRISSSACEDLSAALIANKLITMDLSGNGVGFPMWMLCEGLRH P CKLQMLQKCOLESQACQEWASVLGTPNHLVDLTGNALDEGLRLCCGLRHPC R LRTLWLKICRLTAACBELASTLSVNQSLRELDLSLNFNDLGDWGLWLLAEGLQHPAC RLQTLRLGICRLGSAACEGLSVLQANHLRELDLSFNDLGDWGLWLLAEGLQHPAC RLQTLWDSCLGTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHGPKCLR VLWLFGMDLNMWTHSRLLAALRVTKPYLDIGC"

ORIGIN

ORIGIN

[illegible]

Qy	61	GAGCGTGTGGAAC	TGAAAG	TTCAAG	TTATAC	CTTGGGGA	CCGACACAG	CTGGGAGAA	120
Db	281	GAGCGTGTGGAAC	CTGAAGAAG	TTCAAG	TTATAC	CTTGGGGA	CCGACACAG	CTGGGAGAA	340
Qy	121	GGCAAGATCCCT	TGGGGAAG	CTGGAAG	AGCGCGT	CCCTCGAAAT	GGCCAG	CTGCTC	180
Db	341	GGCAAGATCCCT	TGGGGAAG	CTGGAAG	AGCGCGT	CCCTCGAAAT	TGCCAG	CTGCTC	400
Qy	181	ATCACCCATTTG	GGGCCAG	GAGAGG	CGCTTGG	AGTTGGT	CTCTCAG	CACTTTGAG	240
Db	401	ATCACCCATTTG	GGGCCAG	GAGAGG	CGCTTGG	AGTTGGT	CTCTCAG	CACTTTGAG	460
Qy	241	AAACAGGAAGAC	CTGTGGGAG	GAGGAC	AGAGAG	AGCTGTG	AGG-----	-----	288
Db	461	AAACAGGAAGAC	CTGTGGGAG	GAGGAC	AGAGAG	AGCTGTG	AGGATAC	CCCACT	520
Qy	289	-----	-----	-----	-----	-----	-----	-----	288
Db	521	GGTGGCCCGT	CTCACTTGG	GAAC	CAGTCAA	CATGCTT	CTTGG	AACTCTCTCTT	580
Qy	289	-----	GATCC	CAGGAA	CCCTAC	AGGACTAT	TCTCG	CAGGAAATTCGG	339
Db	581	CCAAGAAAGAT	CCCCAGGAA	ACCTAC	AGGACTAT	TCTCG	CAGGAAATTC	CGGCTCATG	640
Qy	340	GAAGACCGCAAT	GGCGCCTAG	GGGAATGT	CAACCTC	CAGGCC	ACCGGTAC	ACCGGCTC	399
Db	641	GAAGACCGCAAT	GGCGCCTAG	GGGAATGT	GTGTCA	CCCTCAG	CCACCGGT	TACACCGGCTC	700
Qy	400	CTGCTGTGTAAG	GAGACATCAA	ACCCAT	TCGAGT	TCAGG	AGCTTCAG	AGAGCTTC	459
Db	701	CTGCTGTGTAAG	GAGACATCAA	ACCCAT	TCGAGT	TCAGG	AGCTTCAG	AGAGCTTC	760
Qy	460	CGGGACAC	CGGAGGAC	CGTGG	CACAC	CCAGGCTAG	CCCCATCA	AGATAGAG	519
Db	761	CGGGACAC	CGGAGGAC	CGTGG	CACAC	CCAGGCTAG	CCCCATCA	AGATAGAG	820
Qy	520	GAGCCAGAC	GAGGAGCG	CCCCG	AGCCAC	CCGCTG	TATGCA	AGCGCGG	579
Db	821	GAGCCAGAC	GAGGAGCG	CCCCG	AGCCAC	CCGCTG	TATGCA	AGCGCGG	880
Qy	580	ATAGGCAAGTC	CATGCTGG	CACAC	AGGTGAT	CTGG	ACTGGG	CACCGG	639
Db	881	ATAGGCAAGTC	CATGCTGG	CACAC	AGGTGAT	CTGG	ACTGGG	CACCGG	940
Qy	640	CAAGCCAGATTT	GAATTATCT	CTTCTAC	ATCAA	CTGCAGG	GAGATGA	ACCAGAGT	699
Db	941	CAAGCCAGATTT	GAATTATCT	CTTCTAC	ATCAA	CTGCAGG	GAGATGA	ACCAGAGT	1000
Qy	700	GAATGACAGATG	CAAGA	CCCTCAT	CTTCTCAG	CTGCTGG	CCAGCC	AGCGGCTCT	759
Db	1001	GAATGACAGATG	CAAGA	CCCTCAT	CTTCTCAG	CTGCTGG	CCAGCC	AGCGGCTCT	1060
Qy	760	GAGCTCAT	CGAGTTCC	CGAGCG	CTCTTTT	CATCAT	CGA	CGGCTTCG	819
Db	1061	GAGCTCAT	CGAGTTCC	CGAGCG	CTCTTTT	CATCAT	CGA	CGGCTTCG	1120
Qy	820	CTTTCTTTT	CCACGAT	CTCTCAGG	ACCTGT	GTGCTGT	GGGAGG	AGAAAC	879
Db	1121	CTTTCTTTT	CCACGAT	CTCTCAGG	ACCTGT	GTGCTGT	GGGAGG	AGAAAC	1180
Qy	880	GAGCTGCTTT	TACAGCTTA	TTTCG	GAAGA	AGAGT	GTCTCC	TGAGCTAT	939
Db	1181	GAGCTGCTTT	TACAGCTTA	TTTCG	GAAGA	AGAGT	GTCTCC	TGAGCTAT	1240
Qy	940	ACCACAGG	CCCCAGGCTTT	GGGAA	AGCTCC	ACCGTCTG	CTTGG	AGCAC	999
Db	1241	ACCACAGG	CCCCAGGCTTT	GGGAA	AGCTCC	ACCGTCTG	CTTGG	AGCAC	1300
Qy	1000	GAGATCT	GTGGGCTTCT	CTGAGG	CAGAA	AGGAAT	ACTTCT	TACAGTAT	1059
Db	1301	GAGATCT	GTGGGCTTCT	CTGAGG	CAGAA	AGGAAT	ACTTCT	TACAGTAT	1360
Qy	1060	GCAGAG	CAGGCGGG	CCAA	GTCTTCAA	TTACGT	GAGGAC	ACAGG	1119

Db	1165	 ACCACGGCCACCGCTTTTGGAGAGCTCCACCGTCTGTGGAGCACCCGACATGTG	1224
Qy	1000	GAGATCCTGGGCTTCTCTGAGGCGAAGAGGATACCTTCTACAGTATTTCCCAAT	1059
Db	1225	GAGATCCTGGGCTTCTCTGAGGCGAAGAGGATACCTTCTACAGTATTTCCCAAT	1284
Qy	1060	GCAGCAGGCGGGCCCAAGTCTTCAAATTAGTGAGGACAAACAGCCCTCTCTCCACATG	1119
Db	1285	GCAGCAGGCGGGCCCAAGTCTTCAAATTAGTGAGGACAAACAGCCCTCTCTCCACATG	1344
Qy	1120	TGCTTCTGCTCCCTGGTGCTGGGTGGTGATACCTGCTCCAGCAGCAGCTGGAGGGT	1179
Db	1345	TGCTTCTGCTCCCTGGTGCTGGGTGGTGATACCTGCTCCAGCAGCAGCTGGAGGGT	1404
Qy	1180	GGGGGGCTGTGACACAGAGCTCCAGACACCACTGCAGTGATACGTCTTACCTGCTG	1239
Db	1405	GGGGGGCTGTGACACAGAGCTCCAGACACCACTGCAGTGATACGTCTTACCTGCTG	1464
Qy	1240	AGTCTGATGAACCCAAAGCCGGGGCCCCCGCTCCAGCCCCCACAACCCAGAGAGGG	1299
Db	1465	AGTCTGATGAACCCAAAGCCGGGGCCCCCGCTCCAGCCCCCACAACCCAGAGAGGG	1524
Qy	1300	TTTGTCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAAATCTTATTGAGGAGCAGGAC	1359
Db	1525	TTTGTCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAAATCTTATTGAGGAGCAGGAC	1584
Qy	1360	CTCCGGAAGCAGCGCTAGACGGGGAAGAGTCTCTGCTCTCCCTCAACATGAACATCTTC	1419
Db	1585	CTCCGGAAGCAGCGCTAGACGGGGAAGAGTCTCTGCTCTCCCTCAACATGAACATCTTC	1644
Qy	1420	CAGAAGACATCAACTGTGAGAGCTACTACAGCTTTCATCCACTTGAGTTTCCAGGAATTC	1479
Db	1645	CAGAAGACATCAACTGTGAGAGCTACTACAGCTTTCATCCACTTGAGTTTCCAGGAATTC	1704
Qy	1480	TTTGACGATGTACTATATCTTGAAGAGGAGGCGGGGAGGCGCCAGACAGGAC	1539
Db	1705	TTTGACGATGTACTATATCTTGAAGAGGAGGCGGGGAGGCGCCAGACAGGAC	1764
Qy	1540	GTGACAGGCTGTGACCGAGTACGCTTTTCTGAAGAGGAGCTTCTCGGCACTCACCAGC	1599
Db	1765	GTGACAGGCTGTGACCGAGTACGCTTTTCTGAAGAGGAGCTTCTCGGCACTCACCAGC	1824
Qy	1600	CGCTTCTGTTGGACTCTCTGAAACGAGGAGCAGGAGCCACCTGGAGAGAGTCTCTGC	1659
Db	1825	CGCTTCTGTTGGACTCTCTGAAACGAGGAGCAGGAGCCACCTGGAGAGAGTCTCTGC	1884
Qy	1660	TGGAAGGCTCTCGCGCACATCAAGATGGAACCTGTTGACAGTGGATCCAAAGCAAGCTCAG	1719
Db	1885	TGGAAGGCTCTCGCGCACATCAAGATGGAACCTGTTGACAGTGGATCCAAAGCAAGCTCAG	1944
Qy	1720	AGCAGCGCTCCACCTGCAGAGGCTCTTGGAGTCTTTCAGCTGCTGTGTACAGATC	1779
Db	1945	AGCAGCGCTCCACCTGCAGAGGCTCTTGGAGTCTTTCAGCTGCTGTGTACAGATC	2004
Qy	1780	CAGAGGAGGAGTTTATCCAGAGCCCTGAGCCACTTCCAGGTGATCGTGGTCAGCAAC	1839
Db	2005	CAGAGGAGGAGTTTATCCAGAGCCCTGAGCCACTTCCAGGTGATCGTGGTCAGCAAC	2064
Qy	1840	ATTGCCCTCAAGATGGAGCAGATGCTCTCGTTCTGTCTGAAGCGCTGCAGGAGCGCC	1899
Db	2065	ATTGCCCTCAAGATGGAGCAGATGCTCTCGTTCTGTCTGAAGCGCTGCAGGAGCGCC	2124
Qy	1900	CAGTGTGCTACTGTATGTCGCGCACCTACAGCGGAGCAGGGAGAGCCGCGGAGTGC	1959
Db	2125	CAGTGTGCTACTGTATGTCGCGCACCTACAGCGGAGCAGGGAGAGCCGCGGAGTGC	2184
Qy	1960	TCCGAGGAGCGCACACGCTGTGTGTGAGCTCAGACACAGAGAGGACCGTTCTGCTGGAC	2019
Db	2185	TCCGAGGAGCGCACACGCTGTGTGTGAGCTCAGACACAGAGAGGACCGTTCTGCTGGAC	2241
Qy	2020	GCCTACAGTGAACATCTGGCAGCGGCCCTGTGSCAACATCCAAACCTGATAGAGCTGCT	2079
Db	3151	GGCTGCTGA	3159

Db 1465 AGTCTGATGCAACCAAGCCGGGGGCCCCCGGCTTCAGCCGCCACCAACCAAGAGGG 1524
Qy 1300 TTGTGCTCTTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1359
Db 1525 TTGTGCTCTTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1584
Qy 1360 CTCGGAAGCAGCGCTAGACGGGAGAGAGCTCTGCTGCTTCTTCAACATGAACATCTTC 1419
Db 1585 CTCGGAAGCAGCGCTAGACGGGAGAGAGCTCTGCTGCTTCTTCAACATGAACATCTTC 1644
Qy 1420 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTATCCACTTGAGTTTCCAGGAATTC 1479
Db 1645 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTATCCACTTGAGTTTCCAGGAATTC 1704
Qy 1480 TTTCAGCTATGTACTATATCTTGAGAGGGGAGGCGGGGAGGCCCCCAGACAGGAC 1539
Db 1705 TTTCAGCTATGTACTATATCTTGAGAGGGGAGGCGGGGAGGCCCCCAGACAGGAC 1764
Qy 1540 GTGACCAAGGCTTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTGGCACTCACAGC 1599
Db 1765 GTGACCAAGGCTTTGACCGAGTACGCGTTTCTGAAAGGAGCTTCTGGCACTCACAGC 1824
Qy 1600 CGCTTCTGTTTGGACTCTCTGAAAGGAGACAGGAGCCACTGGAGAGAGTCTTGC 1659
Db 1825 CGCTTCTGTTTGGACTCTCTGAAAGGAGACAGGAGCCACTGGAGAGAGTCTTGC 1884
Qy 1660 TGGAGGCTCTGCGCGCATCAAGATGGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 1719
Db 1885 TGGAGGCTCTGCGCGCATCAAGATGGACCTGTTGCAAGTGGATCCAAAGCAAGCTCAG 1944
Qy 1720 AGCGAGGCTCCAGCGCTGACAGGCTCTGAGTCTTCAAGTCTTCAAGTCTTCAAGATC 1779
Db 1945 AGCGAGGCTCCAGCGCTGACAGGCTCTGAGTCTTCAAGTCTTCAAGTCTTCAAGATC 2004
Qy 1780 CAGGAGGAGGAGTTTATCCAGAGGCTGAGGACCTTCCAGGATGATCGTGGTCAAGAAC 1839
Db 2005 CAGGAGGAGGAGTTTATCCAGAGGCTGAGGACCTTCCAGGATGATCGTGGTCAAGAAC 2064
Qy 1840 ATTGCTCCAGATGGAGCAATGGTCTCTGCTTCTGTTGAAAGCGCTGAGAGGCGCC 1899
Db 2065 ATTGCTCCAGATGGAGCAATGGTCTCTGCTTCTGTTGAAAGCGCTGAGAGGCGCC 2124
Qy 1900 CAGTGTGCTGCTTGTATGGCGCACCTTACAGCGGAGCGGAGAGCGCGGAGGTC 1959
Db 2125 CAGTGTGCTGCTTGTATGGCGCACCTTACAGCGGAGCGGAGAGCGCGGAGGTC 2184
Qy 1960 TCCGAGAGGCGCACAGCTGTTGTGAGCTCAGACAGAGAGGACCGTCTGCTGGAC 2019
Db 2185 TCCGAGAGGCGCACAGCTGTTGTGAGCT--ACAGAGAGGACCGTCTGCTGGAC 2241
Qy 2020 GCCTACAGTGAACATCTGGCAGCGGCTCTGTGACCAATCCAACTGATAGAGTGTCT 2079
Db 2242 GCCTACAGTGAACATCTGGCAGCGGCTCTGTGACCAATCCAACTGATAGAGTGTCT 2301
Qy 2080 CTGTACCGAATGCCCTGGGAGCGGCGGCTGAGCTGCTGTGCAAGGACTCAGAC 2139
Db 2302 CTGTACCGAATGCCCTGGGAGCGGCGGCTGAGCTGCTGTGCAAGGACTCAGAC 2361
Qy 2140 CCCAACTGCAAACTTCAGAACTCAGGCTGAAGAGGTCGCGCATCTCCAGCTCAGCCTGC 2199
Db 2362 CCCAACTGCAAACTTCAGAACTCAGGCTGAAGAGGTCGCGCATCTCCAGCTCAGCCTGC 2421
Qy 2200 GAGGAGCTCTCTGAGCTCTCATGACCAATTAAGAAATTTGACAAAGATGGATCTCAGTGGC 2259
Db 2422 GAGGAGCTCTCTGAGCTCTCATGACCAATTAAGAAATTTGACAAAGATGGATCTCAGTGGC 2481
Qy 2260 AACGGCTGTTGATTTCCAGGATGATGCTGCTTTGCGAGGCGCTCGCGCATCCCGAGTGC 2319
Db 2482 AACGGCTGTTGATTTCCAGGATGATGCTGCTTTGCGAGGCGCTCGCGCATCCCGAGTGC 2541
Qy 2320 AGGCTGAGATGATTTCAAGTTGAGGAAGTGTGAGCTGAGTCCGGGCTTGTGAGGAGATG 2379
Db 2542 AGGCTGAGATGATTTCAAGTTGAGGAAGTGTGAGCTGAGTCCGGGCTTGTGAGGAGATG 2601

Qy 2380 GCTTCTGTCTGGGACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2439
Db 2602 GCTTCTGTCTGGGACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2661
Qy 2440 GAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGACCCAGTCTGAGACTACGG 2499
Db 2662 GAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGACCCAGTCTGAGACTACGG 2721
Qy 2500 ACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCTGACGAGCTGGCCTCAACT 2559
Db 2722 ACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCTGACGAGCTGGCCTCAACT 2781
Qy 2560 CTCAGTGTGAACCAAGAGCTGAGAGAGCTGAGACTGAGCTGAGCTGAGGAGCCTC 2619
Db 2782 CTCAGTGTGAACCAAGAGCTGAGAGAGCTGAGACTGAGCTGAGCTGAGGAGCCTC 2841
Qy 2620 GGGGTGCTGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCCTC 2679
Db 2842 GGGGTGCTGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCCTC 2899
Qy 2680 TTGGGCACTCTGCGGCTGGGCTCTGCGGCTGCTGAGGAGCTTCTGCTGCTGCTCAGGCT 2739
Db 2900 ----- 2899
Qy 2740 AACCAACCTCCGGGAGCTGGAATTGAGTTTCAACGAGACTGGGAGACTGGGAGCTGTGG 2799
Db 2900 ----- 2899
Qy 2800 TTGCTGCTGAGGGCTGCAACATCCCGCTGAGACTCCAGAACTGCTGGTGGATAGC 2859
Db 2900 ----- 2910
Qy 2860 TGTGCTCAGACCAAGGCTTGTGAGAACTTTACTTCACTGGGATCAACAGGAC 2919
Db 2911 TGTGCTCAGACCAAGGCTTGTGAGAACTTTACTTCACTGGGATCAACAGGAC 2970
Qy 2920 TTGACGAGCTTTACTGACCAACAAAGCTGAGGAGCAGAGTGTCCGACTGCTTGC 2979
Db 2971 TTGACGAGCTTTACTGACCAACAAAGCTGAGGAGCAGAGTGTCCGACTGCTTGC 3030
Qy 2980 AAGCGCTGAGCCATCTGGCTGCAAACTCGAGTCTCTGTTATTTGGGATGAGCTG 3039
Db 3031 AAGCGCTGAGCCATCTGGCTGCAAACTCGAGTCTCTGTTATTTGGGATGAGCTG 3090
Qy 3040 AATAAATGACCCAGTAGTGTGGCAGCTTCCAGTAAACAAACCTTTATTTGAGACTT 3099
Db 3091 AATAAATGACCCAGTAGTGTGGCAGCTTCCAGTAAACAAACCTTTATTTGAGACTT 3150
Qy 3100 GGCTGCTGA 3108
Db 3151 GGCTGCTGA 3159

RESULT 10
AY116206

LOCUS AY116206 3395 bp mRNA linear PRI 21-MAY-2003
DEFINITION Homo sapiens monarch-1 splice form III mRNA, complete cds;
alternatively spliced.

ACCESSION AY116206
VERSION AY116206.1 GI:21711824

KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3395)

AUTHORS Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.

TITLE Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat

Protein That Controls Classical and Nonclassical MHC Class I Genes

J. Immunol. 170 (11), 5354-5358 (2003)

PUBMED 12759408

REFERENCE 2 (bases 1 to 3395)

AUTHORS
TITLE
JOURNAL

Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
Direct Submission
Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA

FEATURES

source Location/Qualifiers

1. 3395
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221..3073
/notes="contains alternatively spliced exons 7 and 8"
/codon_start=1
/product="monarch-1 splice form III"
/protein_id="AAW75144.1"
/db_xref="GI:21711825"
/translation="MURTAGRDLCLSTVBLEAVELEKFKLYLGTATELGEGBKIP
WGMKAGLEMAQLLITHPGEEAWRLALSTFERINRDLWREGQEDLVRDTPPG
PSLGNQSTCLLEVLITPRKDPQETRYDRVRKFRIMEDNRALGECVNLSHRYRL
LLVKGHSNPMVOQLLDTGRGHARTVGHQASPIKTETLFEDEERPEPRTVMVGA
AGIGKSLAHKVMLDKGLFOGRDYLFYINCREMNQSADECSQDLIFSCWPEPS
APLOELLRPERLLFIIDGDELKPSFDFPGWCLCWEKEKRETELLLSLIRKLLP
EUSLLITPTALEKHLRLEHPRHVEIILGFEAEKKEYFYKYFHNAEQGVFNRYR
DNEPLFTMCVPLVCCVCTCQQQLGEGGLRQTSRTTAVVYMLYLLSLMOPKPGAP
RLOPPNQRGLCSLAADYLNOKILFEEQDLRKHGDDVSAFLNMNI FOKDINER
YYSFHLSPQEFPAAMYIYLDEGEGAGDPQDVTRLLTEVAPSERFLALSRFLGL
LNEETRSHEKSLCWKVSPIKMDLLQWIOSKAQSDGSLTQQGSLBFFSLBEIREE
FTQALSHQVIVSNIASGHEMWSFCLKRCRQAQVHLHYGATYSADGEDRACSLA
KQHLVLRPERTVLLDAYSEHLAAALCTNPNI LELSYRNALSGRKYLLLCQGLRH
PNCQNLRLKRCRISSACEDLSAALIANKNLTRMDLNGVGFPMMLLCEGLRHP
QCRLOMIQRLKCOLESQAQEMASVLGTPNHLVELDLTGNALEDLGLRLCCGLRHPV
CRLRTLWLDSCGLTAKACENLYFTLGINOTLTDLYLTNNALDGTGVRLCKRLSHP
GCKRLVWLFGMDLKNMTHSRLLAALRVTPYLDIGC"
869..892
/notes="Region: P-loop"
2069..3070
/notes="Region: leucine rich repeat"
221..289
/notes="Region: pyrin domain"

misc_feature

misc_feature

misc_feature

ORIGIN

Query Match 79.5%; Score 2470; DB 9; Length 3395;
Best Local Similarity 94.5%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 65; Indels 87; Gaps 2;

QY 1 ATGCTACGAAACCGCAGGAGGAGCGCCCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
DB 221 ATGCTACGAAACCGCAGGAGGAGCGCCCTGTGCGCTGTCCACCTACTTGGAGAACTC 280

QY 61 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCAGACAGAGCTGGGAGAA 120
DB 281 GAGGCTGTGGAAGTCAAGATTATACCTGGGACCGCAGACAGAGCTGGGAGAA 340

QY 121 GCGAAGATCCCTGGGAGACATGGAAGGCGGTCCCTTGGAAATGGCCAGCTGTCTC 180
DB 341 GCGAAGATCCCTGGGAGACATGGAAGGCGGTCCCTTGGAAATGGCCAGCTGTCTC 400

QY 181 ATCAACCATTCGGGACAGAGGCTGTGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
DB 401 ATCAACCATTCGGGACAGAGGCTGTGAGTTGGCTCTCAGACCTTTGAGCGGATA 460

QY 241 AACAGGAAGGACTGTGGGAGAGAGCAGAGAGGACCTGTGAGG----- 288
DB 461 AACAGGAAGGACTGTGGGAGAGAGCAGAGAGGACCTGTGAGGATACCCACCT 520

QY 289 ----- 288
DB 521 GGTGGCCGCTCCTCACTTGGGAACCACTCAATGCTTCTTGAAGTCTCTCTTGTCACT 580

QY 289 -----GATCCCAGGAACCTACAGGACTATGTCCGAGGAAATTCGGGCTCATG 339
DB 581 CCAAGAAAAGATCCCCAGGAAACCTACAGGACTATGTCCGAGGAAATTCGGGCTCATG 640

2020 GCCTACAGTGAACATCTGGCAGCGGCCCTGTGCACCAATCAAACTGATAGAGCTGTCT 2079
|
|
|
2321 GCCTACAGTGAACATCTGGCAGCGGCCCTGTGCACCAATCAAACTGATAGAGCTGTCT 2380
|
|
|
2080 CTGTACCGAATGCCCTGGGAGCGCGGGGGTGAAGCTGTCTCTCAAGGACTCAGACAC 2139
|
|
|
2381 CTGTACCGAATGCCCTGGGAGCGCGGGGGTGAAGCTGTCTCTCAAGGACTCAGACAC 2440
|
|
|
2140 CCACACTGCAAACTTCAAGAACCTGAGCTGAGAGGTGCGCATCTCCAGCTCAGCTGC 2199
|
|
|
2441 CCACACTGCAAACTTCAAGAACCTGAGCTGAGAGGTGCGCATCTCCAGCTCAGCTGC 2500
|
|
|
2200 GAGGACCTCTCTGAGCTCTCATAGCAATCAAGAAATTTGCAAGGATGGATCTCAGTGGC 2259
|
|
|
2501 GAGGACCTCTCTGAGCTCTCATAGCAATCAAGAAATTTGCAAGGATGGATCTCAGTGGC 2560
|
|
|
2260 AACGGGCTTGATTCACGAGCATATGCTGCTTTTGGAGGGCTGCGGCATCCCGAGTGC 2319
|
|
|
2561 AACGGGCTTGATTCACGAGCATATGCTGCTTTTGGAGGGCTGCGGCATCCCGAGTGC 2620
|
|
|
2320 AGGCTGAGAGTATTCAGTTGAGGAAGTGTGAGCTGAGTCCGGGGCTTTGTCAGGAGATG 2379
|
|
|
2621 AGGCTGAGAGTATTCAGTTGAGGAAGTGTGAGCTGAGTCCGGGGCTTTGTCAGGAGATG 2680
|
|
|
2380 GCTTCTGTCTCGGCACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2439
|
|
|
2681 GCTTCTGTCTCGGCACCAACCCACATCTGTTGAGTTGGACCTGACAGGAAATGCACTG 2740
|
|
|
2440 GAGGATTTGGGCTTGAGTTACTATGTCAGGAGTGTGAGGACCCAGTCTGAGACTACGG 2499
|
|
|
2741 GAGGATTTGGGCTTGAGTTACTATGTCAGGAGTGTGAGGACCCAGTCTGAGACTACGG 2800
|
|
|
2500 ACTTTGTGG 2508
|
|
|
2801 ACTTTGTGG 2809
|
|
|
RESULT 12
LOCUS AC008753/3 147330 bp DNA linear PRI 22-MAR-2003
DEFINITION Homo sapiens chromosome 19 clone CTD-3022G6, complete sequence.
ACCESSION AC008753
VERSION AC008753.9 GI:25140998
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
2 (bases 1 to 147330)
DOE Joint Genome Institute.
REFERENCE
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 147330)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (29-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 147330)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (06-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
5 (bases 1 to 147330)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA

6 (bases 1 to 147330)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (22-NOV-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
7 (bases 1 to 147330)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
AUTHORS Direct Submission
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Nov 21, 2002 this sequence version replaced gi:9937751.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.6% of Sequence;
Estimated Total Number of Errors is 0.5.
NOTE: This insert is not the entire sequence of the clone (entire
sequence is 160kb). It is clipped at the overlap with AC008440. The
number of bases overlapped is 22832.
FEATURES
source
1. 147330
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTD-3022G6"
ORIGIN
Query Match 54.8%; Score 1703; DB 9; Length 147330;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1703; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 289 GATCCCCAGGAAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 348
DB 80021 GATCCCCAGGAAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATGGAAGACCGC 79962
QY 349 AATCGGCGCTTAGGGGAATGTCAACTCAGCCACCGGTACACCCGCTCTCTGCTGTG 408
DB 79961 AATCGGCGCTTAGGGGAATGTCAACTCAGCCACCGGTACACCCGCTCTCTGCTGTG 79902
QY 409 AAGGAGCACTCAAAACCCCATCGAGTCCAGCAGCAGCTTCTGGACACAGGCGGGGACAC 468
DB 79901 AAGGAGCACTCAAAACCCCATCGAGTCCAGCAGCAGCTTCTGGACACAGGCGGGGACAC 79842
QY 469 GCGAGGACCGTGGGACACCCAGCTAGCCCCATCAAGATAGAGACCCCTTTTGAAGCAGAC 528
DB 79841 GCGAGGACCGTGGGACACCCAGCTAGCCCCATCAAGATAGAGACCCCTTTTGAAGCAGAC 79782
QY 529 GAGGAGCGCCCGAGGACACCCGCGCACCGTGTCTATGCAAGGCGGGCAGGAGTAGGCAAG 588
DB 79781 GAGGAGCGCCCGAGGACACCCGCGCACCGTGTCTATGCAAGGCGGGCAGGAGTAGGCAAG 79722
QY 589 TCCATGTGGCACACAAGGTGATGCTGGACTGGGGGACCGGAAGCTCTTCCAGGCGACA 648
DB 79721 TCCATGTGGCACACAAGGTGATGCTGGACTGGGGGACCGGAAGCTCTTCCAGGCGACA 79662
QY 649 TTTGATTTATCTTTCTACATCACTGAGGAGATGAACCCAGAGTGCCACCGAATGCAGC 708
DB 79661 TTTGATTTATCTTTCTACATCACTGAGGAGATGAACCCAGAGTGCCACCGAATGCAGC 79602
QY 709 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCCGAGGCGCTCTCCAGGAGCTCATC 768
DB 79601 ATGCAAGACCTCATCTTTCAGCTGTGGCTGAGCCCGAGGCGCTCTCCAGGAGCTCATC 79542
QY 769 CGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGATGAGCTCAAGCTTCTTTTC 828
DB 79541 CGAGTTCCCGAGCGCTCTTTTCATCATCGAGGCTTCGATGAGCTCAAGCTTCTTTTC 79482
QY 829 CACGATCTCCAGGACCTGTGTGCTCTGCTGGAGGAGAAACCGGCCACCGAGCTGCTT 888
DB 79481 CACGATCTCCAGGACCTGTGTGCTCTGCTGGAGGAGAAACCGGCCACCGAGCTGCTT 79422

Db 1314 -----AGGAGATGAACAGAGTGCCACGGAAATGCAGC 1346
Qy 709 ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGAGAGTCAATC 768
Db 1347 ATGCAAGACCTCATCTTCAGCTGTGGCTGAGCCAGCGGCTCTCCAGAGAGTCAATC 1406
Qy 769 CGAGTTCCCGAGCGCTCTCTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCCTTTTC 828
Db 1407 CGAGTTCCCGAGCGCTCTCTTTTCATCATCGACGCTTCGATGAGCTCAAGCTTCCTTTTC 1466
Qy 829 CACGATCCTCAGGAGCCCTGTGCTCTGCTGGAGAGAAACGGCCACCGAGCTGCTTT 888
Db 1467 CACGATCCTCAGGAGCCCTGTGCTCTGCTGGAGAGAAACGGCCACCGAGCTGCTTT 1526
Qy 889 CTTAAACAGCTTAAATTCGGAAGAGCTGCTCCTGTAGCTATCTTTGCTCATACACACAGG 948
Db 1527 CTTAAACAGCTTAAATTCGGAAGAGCTGCTCCTGTAGCTATCTTTGCTCATACACACAGG 1586
Qy 949 CCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGACCCCGAGCATGTGGAGATCCTG 1008
Db 1587 CCCAGGCTTTGGAGAGCTCCACGCTCTGCTGGAGACCCCGAGCATGTGGAGATCCTG 1646
Qy 1009 GGCCTTCTCTGAGCGAAGAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1068
Db 1647 GGCCTTCTCTGAGCGAAGAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1706
Qy 1069 GCGGCCCAAGTCTTCAATTCAGTGAAGGACAAAGAGCTCTCTTACCAATGTCTTCGTC 1128
Db 1707 GCGGCCCAAGTCTTCAATTCAGTGAAGGACAAAGAGCTCTCTTACCAATGTCTTCGTC 1766
Qy 1129 CCCCTGTGTGCTGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG 1188
Db 1767 CCCCTGTGTGCTGGTGTGTACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGCTG 1826
Qy 1189 TTGAGACAGACCTCCAGGACCACTGCAGTGTACATCTCTACCTGCTGAGTCTGATG 1248
Db 1827 TTGAGACAGACCTCCAGGACCACTGCAGTGTACATCTCTACCTGCTGAGTCTGATG 1886
Qy 1249 CAAACCAAGCGGGGCGCCCGGCTCCAGCCGCCACCAACAGAGAGGGTTGTCTCC 1308
Db 1887 CAAACCAAGCGGGGCGCCCGGCTCCAGCCGCCACCAACAGAGAGGGTTGTCTCC 1946
Qy 1309 TTGGCGGAGATGGGCTCTGGAAATCAGAAATCTCTATTTGAGAGCAGGACCTCCGGAAG 1368
Db 1947 TTGGCGGAGATGGGCTCTGGAAATCAGAAATCTCTATTTGAGAGCAGGACCTCCGGAAG 2006
Qy 1369 CACGGCTAGACGGGGAAGACGCTCTGCTCTCCATCAACATGAACATCTTCCAGAGGAC 1428
Db 2007 CACGGCTAGACGGGGAAGACGCTCTGCTCTCCATCAACATGAACATCTTCCAGAGGAC 2066
Qy 1429 ATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGTTTCCAGGAATCTTTGAGCT 1488
Db 2067 ATCAACTGTGAG----- 2078
Qy 1489 ATGTAATAATCCTGACGAGGGGAGGGCGGCGAGCCACAGGACGCTGACCAGG 1548
Db 2079 ----- 2078
Qy 1549 CTGTTGACCGAGTACGCGTTTCTGAAAGAGCTTCTGTGCACTCACCGCCCTTCTCTG 1608
Db 2079 -----AGGAGCTTCTGCACTCACCGCCCTTCTCTG 2111
Qy 1609 TTTGGACTCTGAAACGAGAGACACGAGCCACTGAGAGAGTCTCTGCTGGAGGTC 1668
Db 2112 TTTGGACTCTGAAACGAGAGACACGAGCCACTGAGAGAGTCTCTGCTGGAGGTC 2171
Qy 1669 TCGCCGCAATCAAGATGACCTGTTGCACTGATCCAAAGCAAGACTCAGACGAGCGC 1728
Db 2172 TCGCCGCAATCAAGATGACCTGTTGCACTGATCCAAAGCAAGACTCAGACGAGCGC 2231
Qy 1729 TCCACCTCGACGAGGCTCTCTTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG 1788

Db 2232 TCCACCTCGACGAGGCTCTCTTGGAGTCTTTCAGCTGCTTTACGAGATCCAGGAGGAG 2291
Qy 1789 GAGTTTATCCAGCAGGCTCTGAGCACCTTCAGGTTGATCGTGGTCAGCAACATTCCTCTCC 1848
Db 2292 GAGTTTATCCAGCAGGCTCTGAGCACCTTCAGGTTGATCGTGGTCAGCAACATTCCTCTCC 2351
Qy 1849 AAGATGAGACATAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
Db 2352 AAGATGAGACATAGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2411
Qy 1909 CACTGTTATGGGCAACCTACAGCGCGACGGGGAGACCGCGGAGGTGCTCCGACAGA 1968
Db 2412 CACTGTTATGGGCAACCTACAGCGCGACGGGGAGACCGCGGAGGTGCTCCGACAGA 2471
Qy 1969 GCGCACACGCTGTTGGTGCAGCTCAGACACAGAGGACCGTTCTGCTGCGACCTACAGT 2028
Db 2472 GCGCACACGCTGTTGGTGCAGCT---ACGAGAGGACCGTTCTGCTGCGACCTACAGT 2528
Qy 2029 GAAACATCTGGCAGCGGCTCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGA 2088
Db 2529 GAAACATCTGGCAGCGGCTCTGTGCAACCAATCCAAACCTGATAGAGCTGTCTCTGTACCGA 2588
Qy 2089 AATGCCCTGGGACCGCGGGGTGAAGCTGCTCTGTCAGGACTCAGACACCCCACTGC 2148
Db 2589 AATGCCCTGGGACCGCGGGGTGAAGCTGCTCTGTCAGGACTCAGACACCCCACTGC 2648
Qy 2149 AAACCTTCAGAACCTCA----- 2164
Db 2649 AAACCTTCAGAACCTGAGTAAATTTATCATATATATATATATATTTGAAATATATATA 2708
Qy 2165 ----- 2164
Db 2709 TTGGCCAGGTATGATGGCTCAGCGCTGTAAATTCACAGCACTTTTGGAGGCGCAGATGGGGA 2768
Qy 2165 ----- 2164
Db 2769 GGATCACTTGACCCAGAGTTTCAAGACCAAGCTTGCCCAACATGGTGAACCCCACTCTA 2828
Qy 2165 ----- 2164
Db 2829 CTAAATAATACAAATGAGCAGGCGATGGTGGCACACGCTCTGTAGCCCGCTACTCAGG 2888
Qy 2165 -----GGCTGAAGAGGTGCC 2179
Db 2889 AGGCCAACGCGCAGAGATTGCTTCAACCCAGGAGGAGGTGCTGGCTGAAGAGGTGCC 2948
Qy 2180 GATCTCTCAGCTCAGCTCGAGGAGCTCTCTGAGCTCTCATAGCAATAGAAATTTGA 2239
Db 2949 GCATCTCCAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCAATAGAAATTTGA 3008
Qy 2240 CAAGGATGGATCTCAGTGGCAACGGGTTGGATTCCAGGCGATGCTGCTTGGGAGG 2299
Db 3009 CAAGGATGGATCTCAGTGGCAACGGGTTGGATTCCAGGCGATGATGCTGCTTGGGAGG 3068
Qy 2300 GCTCGGCGATCCAGTGCAGGCTGCAGATGATTCAGTTGAGGAAAGTGTCACTGCTGAGT 2359
Db 3069 GCTCGGCGATCCCAATGAGGCTGCAGATGATTCAGTTGAGGAAAGTGTCACTGCTGAGT 3128
Qy 2360 CCGGGCTTGTGAGGATGGCTTCTGTGCTCGGCCAACCCACATCTGTTGAGTGG 2419
Db 3129 CCGGGCTTGTGAGGATGGCTTCTGTGCTTGGCAACCAACCCACATCTGTTGAGTGG 3188
Qy 2420 ACTGACAGAAATGACCTGGAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTCAGGC 2479
Db 3189 ACTGACAGAAATGACCTGGAGGATTTGGGCTGAGGTTACTATGCCAGGAGCTCAGGC 3248
Qy 2480 ACCAGCTCTGCAGACTTACGGAATTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCCT 2539
Db 3249 ACCAGCTCTGCAGACTTACGGAATTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCCT 3308
Qy 2540 GTGACAGCTGGCCTCAACTCTCAGTGTGAAACAGAGCCTGAGAGCTGGACCTGAGCC 2599
Db 3309 GTGACAGCTGGCCTCAACTCTCAGTGTGAAACAGAGCCTGAGAGCTGGACCTGAGCC 3368

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 00:14:00 ; Search time 1131 Seconds
(without alignments)
11674.093 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctagacacgcaggcag.....attggacatggctgctga 3108

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002s:*

7: geneseqn2003as:*

8: geneseqn2003bs:*

9: geneseqn2003cs:*

10: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	3108	100.0	3108	6	ADE36416	Ade36416 Human PAA
2	3057	98.4	3300	6	AAL47129	Aal47129 Pyrin dom
3	2997.8	96.5	3186	6	AAL44363	Aal44363 Human PYR
4	2449.4	78.8	3306	9	ADC30316	Adc30316 Human nov
5	1705.1	54.9	1800	7	ADC03623	Adc03623 Novel hum
6	1681.8	54.1	2158	6	AB578719	Ab578719 Human cDN
7	1616	52.0	4931	6	ABL59333	Ab159333 Nucleotid
8	1394.8	44.9	1683	7	ADC03624	Adc03624 Novel hum
9	800.4	25.8	3857	4	AAD14323	Aad14323 Human PYR
10	800.4	25.8	3857	8	ABX93556	Abx93556 Huma cDNA
11	800.4	25.8	3857	8	ACD27909	AcD27909 Human PYR
12	639	20.6	2524	6	Ade36451	Ade36451 Human PAA
13	594.6	19.1	2847	4	AA501487	Aa501487 Human sec
14	594.6	19.1	2847	7	AB273494	Ab273494 Secreted
15	594.6	19.1	2847	9	ADA98038	Ada98038 Human sec
16	594.6	19.1	2847	9	ADC20194	Adc20194 Human sec
17	588.4	18.9	591	5	ADC32201	Adc32201 Human nov
18	477.8	15.4	506	5	AAS68757	Aas68757 DNA encod
19	470.4	15.1	1557	6	AAL47143	Aal47143 Pyrin dom
20	453.4	14.6	487	8	ACH36225	Ach36225 Human end
21	450.8	14.5	479	3	AAC76566	Aac76566 Human ORF
22	404.4	13.0	4035	6	AAL44356	Aal44356 Human PYR
23	402.8	13.0	3226	6	ABX97181	Abx97181 Human NOV

ALIGNMENTS

RESULT 1

ADE36416

ID ADE36416 standard; cDNA; 3108 BP.

XX AC ADE36416;

XX AC ADE36416;

DT 29-JAN-2004 (first entry)

XX DE Human PAAD and nucleotide binding protein PAN6 cDNA.

XX KW cytostatic; immunosuppressive; vulnery; antiinflammatory; vasotropic;
 KW antiallergic; antiulcer; dermatological; cerebroprotective; cardiant;
 KW antiparkinsonian; nootropic; neuroprotective; anti-HIV; gene therapy;
 KW NFkappaB activation inhibitor; PAAD domain containing polypeptide;
 KW PAAD and nucleotide binding protein 2-6; PAN 2-6; pyrin 2;
 KW apoptosis-associated speck-like protein; caspase recruitment domain 2;
 KW ASC-2; nuclear factor kappa B activation inhibitor; NB-ARC domain;
 KW apoptosis; NFkappaB induction; cytokine processing;
 KW cytokine receptor; signaling caspase-mediated proteolysis;
 KW c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
 KW inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
 KW neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
 KW fibrosis; smooth muscle cell proliferation; balloon angioplasty;
 KW restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
 KW arthritis; lupus; schrojen's syndrome; Crohn's disease;
 KW ulcerative colitis; graft versus host disease; stroke; heart failure;
 KW neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
 KW cancer therapy; PAAD domain family; human; PAN6; gene; ss.

XX Homo sapiens.

XX US2003077699-A1.

XX 24-APR-2003.

XX 25-SEP-2001; 2001US-00965621.

XX 26-SEP-2000; 2000US-00671760.

XX 26-SEP-2000; 2000US-0367367P.

XX (REED/) REED J C.

XX (GODZ/) GODZIK A.

XX (CHUZ/) CHU Z.

XX (PAWL/) PAWLOWSKI K.

XX (FIOR/) FIORENTINO L.

XX (ARIZ/) ARIZA M E.

Aad41224 Human EMB
 Ada45220 Human MAT
 Ada45218 Human MAT
 Aal47135 Pyrin dom
 Aal47131 Pyrin dom
 Aal47140 Pyrin dom
 Abk48628 Human MAT
 Aad49018 Human MAT
 Aad49018 Human NB-
 Aad02764 Human NAC
 Aad02764 Human NAC
 Acc45143 Human NAC
 Aaf83651 Human CAR
 Abs55497 cDNA enco
 Abs55497 cDNA enco
 Abs56030 cDNA enco
 Acc45151 Human NAC
 Acc45152 Human gas
 Aad12951 Human G-P
 Aad02761 Human NB-
 Aad02762 Human NB-
 Aad02765 Human NAC
 Aal47127 Pyrin dom
 Abq75801 Human MDD

PA (STEHLIK C.
XX Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;
PI Stehlik C;
XX WPI: 2002-471256/50.
DR P-PSDB; ADE34417.
XX Novel isolated PAAD domain containing polypeptide useful for inducing
PT apoptosis by inhibiting nuclear factor kappa B activation and in gene
PR therapy for treating cancer.
XX Claim 2; SEQ ID NO 23; 93pp; English.
XX The invention describes an isolated PAAD domain containing polypeptide
CC (1) comprising 80% identity to the amino acid sequence of PAAD and
CC nucleotide binding protein (PAN) 2-6, p21, p21-2, apoptosis-associated speck
CC -like protein containing a caspase recruitment domain (ASC)-2 fully
CC defined in specification, where (1) is biologically active. (1) is useful
CC for identifying a (1)-associated polypeptide, an agent altering that
CC association and agents that modulate PAAD domain mediated inhibition of
CC nuclear factor kappa B (NF-kappaB). A NB-ARC domain polypeptide is useful
CC for identifying an agent that modulates the activity of the NB-ARC domain
CC of (1). (1) or its functional fragments is useful in altering cellular or
CC biochemical process such as apoptosis, NF-kappaB induction, cytokine
CC processing, cytokine receptor signaling caspase-mediated proteolysis or c
CC Jun N-terminal kinase activation, thus having modulating effect on cell
CC life and death (apoptosis) inflammation, cell adhesion or other cellular
CC or biochemical processes. (1) is useful for treating cancer pathologies,
CC keratinocyte, hyperplasia, neoplasia, keloid benign prostatic
CC hyper trophy, inflammatory hyperplasia, fibrosis, smooth muscle cell
CC proliferation in arteries following balloon angioplasty (restenosis),
CC leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis,
CC lupus, schrofen's syndrome, Crohn's disease and ulcerative colitis, graft
CC versus host disease, stroke, heart failure, neurodegenerative diseases
CC such as parkinson's and Alzheimer's disease, human immunodeficiency virus
CC infection (HIV). (1) is useful for diagnosing cancer or monitoring cancer
CC therapy. This sequence encodes human PAAD and nucleotide binding protein
CC PAN6.
XX
XX Sequence 3108 BP; 682 A; 883 C; 931 G; 612 T; 0 U; 0 Other;
Query Match 100.0%; Score 3108; DB 6; Length 3108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3108; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTACGAAACCGCAGGAGGAGCGCTCTGCTCGCTGCTCCACTTCTGGAGAACTC 60
DB 1 ATGCTACGAAACCGCAGGAGGAGCGCTCTGCTCGCTGCTCCACTTCTGGAGAACTC 60
QY 61 GAGGCTGTGAACTGAAAGAGTTCAAGTTATACCTGGGACCGCAGAGCTGGGAGAA 120
DB 61 GAGGCTGTGAACTGAAAGAGTTCAAGTTATACCTGGGACCGCAGAGCTGGGAGAA 120
QY 121 GGCAGATCCCTGGGAGAGCATGGAGAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
DB 121 GGCAGATCCCTGGGAGAGCATGGAGAGCGGCTCCCTGGAAATGGCCAGCTGCTC 180
QY 181 ATCAACCCACTTCGGGCGCAGAGGAGGCTTGGAGTTGGCTCTCAGACCTTTTGACGGATA 240
DB 181 ATCAACCCACTTCGGGCGCAGAGGAGGCTTGGAGTTGGCTCTCAGACCTTTTGACGGATA 240
QY 241 AACAGNAGACCTGTGGGAGAGAGGACAGAGAGACCTGTGGAGGATCCCGAGGAA 300
DB 241 AACAGNAGACCTGTGGGAGAGAGGACAGAGAGACCTGTGGAGGATCCCGAGGAA 300
QY 301 ACCTACAGGACTATGTCGCGAGAAATTCGGCTCATGAGACCGCAATGGCGCTTA 360
DB 301 ACCTACAGGACTATGTCGCGAGAAATTCGGCTCATGAGACCGCAATGGCGCTTA 360
QY 361 GGGGAATGTCTAACCTCAGCCACCGGTACACCCGGCTCCTGTGTGGAGAGCACTCA 420
DB 361 GGGGAATGTCTAACCTCAGCCACCGGTACACCCGGCTCCTGTGTGGAGAGCACTCA 420

QY 421 AACCCCATGAGTCCAGCAGCAGCTTCTGACACAGGCGCGGACACGCGAGACCGTGT 480
DB 421 AACCCCATGAGTCCAGCAGCAGCTTCTGACACAGGCGCGGACACGCGAGACCGTGT 480
QY 481 GGACACAGGCTGACCCCATCAAGATAGACCTCTTTTGAGCCAGAGAGCGGCC 540
DB 481 GGACACAGGCTGACCCCATCAAGATAGACCTCTTTTGAGCCAGAGAGCGGCC 540
QY 541 GAGCACCAGGCGCAGGCTGATCAAGAGCGGCGGAGGATAGGCAAGTCCATGCTGCA 600
DB 541 GAGCACCAGGCGCAGGCTGATCAAGAGCGGCGGAGGATAGGCAAGTCCATGCTGCA 600
QY 601 CACAAGGTGATGCTGCACTGGGCGGAGGAGGAGCTTCTTCCAGGAGATTTGATTATCTC 660
DB 601 CACAAGGTGATGCTGCACTGGGCGGAGGAGGAGCTTCTTCCAGGAGATTTGATTATCTC 660
QY 661 TTCTACATCAACTGACGAGGAGATGAACAGAGTGCACGGAATGAGCATGCAAGACCTC 720
DB 661 TTCTACATCAACTGACGAGGAGATGAACAGAGTGCACGGAATGAGCATGCAAGACCTC 720
QY 721 ATCTTACAGTCTGCTGAGCGCCAGCGGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
DB 721 ATCTTACAGTCTGCTGAGCGCCAGCGGCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
QY 781 CGCCTCTCTTTCATCATGAGCTGAGTGCATGAGCTCAAGCCTTCTTCCAGGATCTCTCAG 840
DB 781 CGCCTCTCTTTCATCATGAGCTGAGTGCATGAGCTCAAGCCTTCTTCCAGGATCTCTCAG 840
QY 841 GGACCTCTGCTCTGCTGGGAGGAGAAACCGGCCACGAGCTGCTTTTAAACAGCTTA 900
DB 841 GGACCTCTGCTCTGCTGGGAGGAGAAACCGGCCACGAGCTGCTTTTAAACAGCTTA 900
QY 901 ATTCCGAGAGAGCTGCTCCCTGAGCTATCTTGTCTATCACCACACGCGCCACCGCTTG 960
DB 901 ATTCCGAGAGAGCTGCTCCCTGAGCTATCTTGTCTATCACCACACGCGCCACCGCTTG 960
QY 961 GAGAAGCTCCACCGCTGCTGGAGCACCCAGGAGCTGAGAGATCTCTGGGCTTCTCTGAG 1020
DB 961 GAGAAGCTCCACCGCTGCTGGAGCACCCAGGAGCTGAGAGATCTCTGGGCTTCTCTGAG 1020
QY 1021 GCAGAAAGAGGAGTACTTCTACAGTATTTTCAAGTATTTTCAAGTATTTTCAAGTATTT 1080
DB 1021 GCAGAAAGAGGAGTACTTCTACAGTATTTTCAAGTATTTTCAAGTATTTTCAAGTATTT 1080
QY 1081 TTCAATTAAGTGGGAGAACGAGCTCTCTTCAACATGCTGCTGCTCCCTCCCTGGTGTGC 1140
DB 1081 TTCAATTAAGTGGGAGAACGAGCTCTCTTCAACATGCTGCTGCTCCCTCCCTGGTGTGC 1140
QY 1141 TGGGTGCTGTGTACTCTCCAGCAGCAGCTGGAGGCTGGGGGCTGTTGAGACAGACG 1200
DB 1141 TGGGTGCTGTGTACTCTCCAGCAGCAGCTGGAGGCTGGGGGCTGTTGAGACAGACG 1200
QY 1201 TCCAGGACCACTGAGTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
DB 1201 TCCAGGACCACTGAGTGTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
QY 1261 GGGGCCCCGCGCTCCAGCCCCCAGCAGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
DB 1261 GGGGCCCCGCGCTCCAGCCCCCAGCAGAGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
QY 1321 GGGCTCTGGAATCAGAAATCTTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGAC 1380
DB 1321 GGGCTCTGGAATCAGAAATCTTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGAC 1380
QY 1381 GGGGAGAGCTCTCTGCT 1440
DB 1381 GGGGAGAGCTCTCTGCT 1440
QY 1441 AGGTACTACAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1500
DB 1441 AGGTACTACAGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1500

QY 1501 CTGACGAGGGGGAGGGCGGGGAGCCAGACAGAGAGCTGACAGAGCTGTGTACCGAG 1560
DB 1501 CTGACGAGGGGGAGGGCGGGGAGCCAGACAGAGAGCTGACAGAGCTGTGTACCGAG 1560
QY 1561 TACGGGTTTTCTGAAAGAGCTTCTGTCGACCTACACAGCGGCTTCTGTTGGAGCTCTG 1620
DB 1561 TACGGGTTTTCTGAAAGAGCTTCTGTCGACCTACACAGCGGCTTCTGTTGGAGCTCTG 1620
QY 1621 AACGAGGAGACAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGCTCCGCCGACATC 1680
DB 1621 AACGAGGAGACAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGCTCCGCCGACATC 1680
QY 1681 AAGATGACCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGCGGCTCCACCTGCGAG 1740
DB 1681 AAGATGACCTGTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGCGGCTCCACCTGCGAG 1740
QY 1741 CAGGGCTCTTGGAGTTCTTCAAGCTGTTGACGAGTCTTGAAGATCCAGGAGGAGGTTTATCCAG 1800
DB 1741 CAGGGCTCTTGGAGTTCTTCAAGCTGTTGACGAGTCTTGAAGATCCAGGAGGAGGTTTATCCAG 1800
QY 1801 CAGGCTTGAAGCTTCCAGTGTATCTGCTGAGCAATTTGCTTCAAGATGGAGGAC 1860
DB 1801 CAGGCTTGAAGCTTCCAGTGTATCTGCTGAGCAATTTGCTTCAAGATGGAGGAC 1860
QY 1861 ATGGTCTCTGCTTCTGTTGAAGCGCTGAGGAGCGCCAGGCTGCTGCTATGATGCG 1920
DB 1861 ATGGTCTCTGCTTCTGTTGAAGCGCTGAGGAGCGCCAGGCTGCTGCTATGATGCG 1920
QY 1921 GCCACCTACAGCGGGAGGAGACCGCGGAGGTGCTCGGAGGAGCGACACGCTG 1980
DB 1921 GCCACCTACAGCGGGAGGAGACCGCGGAGGTGCTCGGAGGAGCGACACGCTG 1980
QY 1981 TTGTTGACGCTCAGACAGAGAGGACCTTCTGCTGAGCGCTCAGTGAACATCTGGCA 2040
DB 1981 TTGTTGACGCTCAGACAGAGAGGACCTTCTGCTGAGCGCTCAGTGAACATCTGGCA 2040
QY 2041 GCGGCCCTGTGACCAATCCAAACCTGATAGAGTGTCTGTACCGAAATGCCCTGGGC 2100
DB 2041 GCGGCCCTGTGACCAATCCAAACCTGATAGAGTGTCTGTACCGAAATGCCCTGGGC 2100
QY 2101 AGCGGGGGTGAAGCTGCTGTGCAAGGATCTAGACACCCCAACTGCAAACTTCAGAAC 2160
DB 2101 AGCGGGGGTGAAGCTGCTGTGCAAGGATCTAGACACCCCAACTGCAAACTTCAGAAC 2160
QY 2161 CTGAGGCTGAAGAGTGCAGATCTCCAGCTCAGCTCGAGGAGCTCTCTGCGAGCTCTC 2220
DB 2161 CTGAGGCTGAAGAGTGCAGATCTCCAGCTCAGCTCGAGGAGCTCTCTGCGAGCTCTC 2220
QY 2221 ATAGCCAATAAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGTTGGATTTCCAGGC 2280
DB 2221 ATAGCCAATAAGAAATTTGACAAAGATGATCTCAGTGGCAACGGGTTGGATTTCCAGGC 2280
QY 2281 ATGATGCTGCTTTGGAGGGCTGGGGATCCCGAGTCCAGGCTCAGATGATTCAGTTG 2340
DB 2281 ATGATGCTGCTTTGGAGGGCTGGGGATCCCGAGTCCCGAGTCCAGATGATTCAGTTG 2340
QY 2341 AGGAAGTCTCAGCTGAGTCCGGGGCTTGTGAGAGATGGCTTCTGCTCGGACCAAC 2400
DB 2341 AGGAAGTCTCAGCTGAGTCCGGGGCTTGTGAGAGATGGCTTCTGCTCGGACCAAC 2400
QY 2401 CCACATCTGGTTGAGTTGACCTGACAGAAATGACCTGGAGGATTTGGGCTCAGGTTA 2460
DB 2401 CCACATCTGGTTGAGTTGACCTGACAGAAATGACCTGGAGGATTTGGGCTCAGGTTA 2460
QY 2461 CTATGCCAGGAGCTGAGGACCCAGTCTGACAGATGACGGAATTTGTGCTGAAAGTCTGC 2520
DB 2461 CTATGCCAGGAGCTGAGGACCCAGTCTGACAGATGACGGAATTTGTGCTGAAAGTCTGC 2520
QY 2521 GGCCTCACTGCTGCTGTGACAGCTGGCTCAACTCTCAGTGTGAACAGAGCCTG 2580
DB 2521 GGCCTCACTGCTGCTGTGACAGCTGGCTCAACTCTCAGTGTGAACAGAGCCTG 2580
QY 2581 AGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTGTGAG 2640

DB 2581 AGAGAGCTGGACCTGAGCCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTGTGAG 2640
QY 2641 GGCCTCAGGCAATCCACGTCGAAAGCTCAGACCTTCGCGTTCGGGATCTGCCGGCTGGGC 2700
DB 2641 GGCCTCAGGCAATCCACGTCGAAAGCTCAGACCTTCGCGTTCGGGATCTGCCGGCTGGGC 2700
QY 2701 TCTCCGCTCTGAGGGGCTTTCTGTTGGTCTCCAGGCAACCAACCTCCGGGAGCTG 2760
DB 2701 TCTCCGCTCTGAGGGGCTTTCTGTTGGTCTCCAGGCAACCAACCTCCGGGAGCTG 2760
QY 2761 GACTTGAATTTCAACGACTCGGAGCTGGGGCTTGTGTTGCTGAGGGCTGCAAA 2820
DB 2761 GACTTGAATTTCAACGACTCGGAGCTGGGGCTTGTGTTGCTGAGGGCTGCAAA 2820
QY 2821 CATCCGCTCGAGACTCCAGAACTGTGGCTGATAGCTGTGGCTCACAGCCAAAGCT 2880
DB 2821 CATCCGCTCGAGACTCCAGAACTGTGGCTGATAGCTGTGGCTCACAGCCAAAGCT 2880
QY 2881 TGTGAGATCTTTACTTCAACCTCGGGATCAACAGACTTTGACCGCTTTACCTGACC 2940
DB 2881 TGTGAGATCTTTACTTCAACCTCGGGATCAACAGACTTTGACCGCTTTACCTGACC 2940
QY 2941 AACACGCTTAGGGGACACAGGTTCGAGCTGCTTTGCAAGCGGCTGAGCCATCTGGC 3000
DB 2941 AACACGCTTAGGGGACACAGGTTCGAGCTGCTTTGCAAGCGGCTGAGCCATCTGGC 3000
QY 3001 TGCAAACTCCAGTCTCTGTTATTTGGGATGAGACCTGAATAATGACCCAGTAGG 3060
DB 3001 TGCAAACTCCAGTCTCTGTTATTTGGGATGAGACCTGAATAATGACCCAGTAGG 3060
QY 3061 TTGGCAGCGCTTCAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108
DB 3061 TTGGCAGCGCTTCAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108

RESULT 2

AA147129
ID AA147129 standard; DNA; 3300 BP.
XX AC AA147129;
XX AC AA147129;
DT 20-AUG-2002 (first entry)
XX Pyrin domain containing protein NALP3/PY5-hs coding sequence.
DE Pyrin domain; PYD domain; antinflammatory; antiparkinsonian;
XX antiarteriosclerotic; antipsoriasis; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; psoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX Unidentified.
OS Unidentified.
XX WO200240668-A2.
PN 23-MAY-2002.
XX 30-OCT-2001; 2001WO-EP012545.
XX 15-NOV-2000; 2000DE-01056587.
PR 30-NOV-2000; 2000DE-01059595.
XX (APOT-) APOTEC RES & DEV LTD.
XX Tschoep J, Martinon F;
PI WPI; 2002-427093/45.
XX P-PSDB; AA017857.
PT New DNA encoding protein with pyrin domain, useful for treating diseases

involving impaired signal transduction, particularly inflammation, also proteins and antibodies.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Claim 5: Fig 1: 116pp: German.

The present invention relates the DNA and their encoded proteins, where the proteins contain at least one PYD (pyrin) domain. These can be used to treat diseases associated with impaired intracellular signal transduction, particularly inflammation such as psoriasis, arteriosclerosis, bacterial or viral infections (particularly meningitis and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma, sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's and Parkinson's diseases. The present sequence is a coding sequence of the invention.

Sequence 3300 BP: 726 A: 943 C: 979 G: 652 T: 0 U: 0 Other:

Query Match 98.4%; Score 3057; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 0; Indels 0

1	ATGCTAGAAACCGCAGCAGGAGCGGCCTCTGTGCTCCCTGTCCACCTACTTGGAAAGAACTC	60	Qy
1	ATGCTAGAAACCGCAGCAGGAGCGGCCTCTGTGCGCTGTCCACCTACTTGGAAAGAACTC	60	Db
61	GAGCGTGGAACTTGAAGAGTTTCAAAGTTTATACCTTGGGACCGCCACACAGAGCTGGAGAA	120	Qy
61	GAGCGTGGAACTTGAAGAGTTTCAAAGTTTATACCTTGGGACCGCCACACAGAGCTGGAGAA	120	Db
121	GGCAAGATCCCTCGGGGAAGCATGGAGAAGGCCGCTCCCTTGGAATGCCCCAGCTGTCTC	180	Qy
121	GGCAAGATCCCTCGGGGAAGCATGGAGAAGGCCGCTCCCTTGGAATGCCCCAGCTGTCTC	180	Db
181	ATCACCACTTTCGGCCAGAGAGGCTTGGAGTTTGGCTCTCAGCACTTTGAGCGGATA	240	Qy
181	ATCACCACTTTCGGCCAGAGAGGCTTGGAGTTTGGCTCTCAGCACTTTGAGCGGATA	240	Db
241	AACAGGAGGACCTGTGGGAGAGGACACAGAGAGGAGCTGTCGAGGATCCCCAGGAA	300	Qy
241	AACAGGAGGACCTGTGGGAGAGGACACAGAGAGGAGCTGTCGAGGATCCCCAGGAA	300	Db
301	ACCTACAGGGACTATGTCCGCAGGAAATTTCCGGCTCATGGAAGACCGCAATCGCGCCCTA	360	Qy
301	ACCTACAGGGACTATGTCCGCAGGAAATTTCCGGCTCATGGAAGACCGCAATCGCGCCCTA	360	Db
361	GGGGAATGTGTCAACCTCAGCCAACCGGTACACCCGGCTCTGCTGGTGAAGAGCACTCA	420	Qy
361	GGGGAATGTGTCAACCTCAGCCAACCGGTACACCCGGCTCTGCTGGTGAAGAGCACTCA	420	Db
421	AACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACACGCGAGGACCCGTG	480	Qy
421	AACCCCATGCAAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACACGCGAGGACCCGTG	480	Db
481	GGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGACGAGAGAGCCGCC	540	Qy
481	GGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGACGAGAGAGCCGCC	540	Db
541	GAGCCACCGCGCACCGTGGTCAATGCAAGCGCGGACAGGATAGGCAAGTCCATGTGGCA	600	Qy
541	GAGCCACCGCGCACCGTGGTCAATGCAAGCGCGGACAGGATAGGCAAGTCCATGTGGCA	600	Db
601	CACAAGGTGATGCTGGACTGGCGGACGGGAGCTTCTCCAGGAGCAGTTTCATTTATCTC	660	Qy
601	CACAAGGTGATGCTGGACTGGCGGACGGGAGCTTCTCCAGGAGCAGTTTCATTTATCTC	660	Db
661	TTCTTACATCAACTGCAGGGAGATGAACACAGAGTGGCACCGGAATGCAGCATGCAAGCCTC	720	Qy
661	TTCTTACATCAACTGCAGGGAGATGAACACAGAGTGGCACCGGAATGCAGCATGCAAGCCTC	720	Db
721	ATCTTCAGCTGTGCGCTGAGCCAGCGGCTCTCTCAGGAGCTCATTCGAGTTTCCCGAG	780	Qy
721	ATCTTCAGCTGTGCGCTGAGCCAGCGGCTCTCTCAGGAGCTCATTCGAGTTTCCCGAG	780	Db

Db 1861 ATGTCCTCTCGTTCTGTTGAAGCGCTGCGAGCGCCAGGTCGTGCACTTGATGCG 1920
QY 1921 GCCACCTACAGCGGAGCGGGAACACCGCGAGGTGCTCGCAGGAGCGCACAGCTG 1980
Db 1921 GCCACCTACAGCGGAGCGGGAACACCGCGAGGTGCTCGCAGGAGCGCACAGCTG 1980
QY 1981 TTGTGTCAGCTCAGACAGAGAGGACCGTTCTGTGGAGCGCTACAGTGAACATCTGGCA 2040
Db 1981 TTGTGTCAGCTCAGACAGAGAGGACCGTTCTGTGGAGCGCTACAGTGAACATCTGGCA 2040
QY 2041 GCGGCCCTGTGCAACAACTCAAACTGATAGAGCTGTCTCTACCGAAATGCCCTGGGC 2100
Db 2041 GCGGCCCTGTGCAACAACTCAAACTGATAGAGCTGTCTCTACCGAAATGCCCTGGGC 2100
QY 2101 AGCCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCACTGCAGAACTTCAGAAC 2160
Db 2101 AGCCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCACTGCAGAACTTCAGAAC 2160
QY 2161 CTGAGGCTGAAGAGTGGCGCATCTCCAGCTCAGCTCGGAGGACTCTCTGAGCTCTC 2220
Db 2161 CTGAGGCTGAAGAGTGGCGCATCTCCAGCTCAGCTCGGAGGACTCTCTGAGCTCTC 2220
QY 2221 ATAGCCAAATAGAAATTCACAAAGATGATCTCAGTGGCAACCGGCTTGGATTCCAGGC 2280
Db 2221 ATAGCCAAATAGAAATTCACAAAGATGATCTCAGTGGCAACCGGCTTGGATTCCAGGC 2280
QY 2281 ATGATGCTGCTTTGGAGGGGCTGCTGTGAGGAGTGGCTTCTGAGGAGTGGCTTCTG 2340
Db 2281 ATGATGCTGCTTTGGAGGGGCTGCTGTGAGGAGTGGCTTCTGAGGAGTGGCTTCTG 2340
QY 2341 AGGAAGTGTGAGTGGAGTGGCGGCTTGTGAGGAGTGGCTTCTGAGGAGTGGCTTCTG 2400
Db 2341 AGGAAGTGTGAGTGGAGTGGCGGCTTGTGAGGAGTGGCTTCTGAGGAGTGGCTTCTG 2400
QY 2401 CCACATCTGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2460
Db 2401 CCACATCTGCTGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2460
QY 2461 CTATCCAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2520
Db 2461 CTATCCAGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2520
QY 2521 GCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 GCGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 AGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2640
Db 2581 AGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2640
QY 2641 GGCCTCAGGATCCACAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2700
Db 2641 GGCCTCAGGATCCACAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2700
QY 2701 TCTGCGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
Db 2701 TCTGCGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
QY 2761 GACTTGAGTTTCAACAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2820
Db 2761 GACTTGAGTTTCAACAGCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 2820
QY 2821 CATCCGCTGAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 2880
Db 2821 CATCCGCTGAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCCAGATCC 2880
QY 2881 TGTGAGATCTTTACTTCACTTCCCTGGGATCAACAGAGCTTCAACAGAGCTTCAACAGAGCTT 2940
Db 2881 TGTGAGATCTTTACTTCACTTCCCTGGGATCAACAGAGCTTCAACAGAGCTTCAACAGAGCTT 2940
QY 2941 AACCAAGCCCTAGGGGACACAGTGTCCGACTGCTTTGCAAGCGGCTGAGCCATCTCGGC 3000

Db 2941 AACCAAGCCCTAGGGGACACAGTGTCCGACTGTTTTCAGAGCGCTGAGCCATCTCGGC 3000
QY 3001 TGCAAACTCCGAGTCTCTGTTTATTTGGGATGGACCTGAATAAATGACCCACAGT 3057
Db 3001 TGCAAACTCCGAGTCTCTGTTTATTTGGGATGGACCTGAATAAATGACCCACAGT 3057
RESULT 3
AAL44363
ID AAL44363 standard; cDNA; 3186 BP.
XX
AC AAL44363;
DT 31-OCT-2002 (first entry)
XX
Human PYRIN-8 cDNA sequence #2.
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory response; inflammatory disorder;
KW immune system disorder; Crohn's disease; multiple sclerosis; cancer;
KW leukaemia; autoimmune disorder; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-8.
XX
Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3186
FT /tag= a
FT /product= "Human PYRIN-8 protein #2"
XX
WO200261049-A2.
PD 08-AUG-2002.
XX
31-JAN-2002; 2002WO-US002967.
XX
31-JAN-2001; 2001US-0265231P.
PR 10-SEP-2001; 2001US-0318645P.
XX
(MILL-) MILLENNIUM PHARM INC.
PA (AMHP) WYETH.
XX
Bertin J, Wang W, Blatcher M;
XX
WPI; 2002-627477/67.
DR P-PSDB; AAO15590.
XX
New PYRIN polypeptides and nucleic acids useful for modulating and
diagnosing stress-related, apoptotic and inflammatory responses, or for
treating inflammatory and immune system disorders, cancers, or
neurological diseases.
XX
Claim 4; Fig 8; 167pp; English.
XX
The invention comprises the amino acid and coding sequences of human
PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
useful for modulating and diagnosing stress-related, apoptotic and
inflammatory responses. The PYRIN protein and DNA sequences are useful
for treating inflammatory disorders and immune system disorders (e.g.
Crohn's disease, reactive arthritis, multiple sclerosis, contact
dermatitis, psoriasis, graft rejection, allergies, viral infections and
bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
(e.g. systemic lupus erythematosus and arthritis); and neurological
diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
protein and DNA sequences may also be used in screening assays, detection
assays (e.g. chromosomal mapping, tissue typing or forensic biology),
predictive medicine (e.g. diagnostic assays, clinical trials and
pharmacogenomics) and transcription profiling. The present DNA sequence
encodes a human PYRIN-8 protein
XX
Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other; SQ

Query Match 96.5%; Score 2997.8; DB 6; Length 3186;
Best local Similarity 97.3%; Pred. No. 0;
Matches 3103; Conservative 0; Mismatches 2; Indels 84; Gaps 2;

QY 1 ATGCTAGCAACCGCAGGAGGAGCGGCTCTGTGCGCTGTGCTACCTACTTGGAGAACTC 60
DB 1 ATGCTAGCAACCGCAGGAGGAGCGGCTCTGTGCGCTGTGCTACCTACTTGGAGAACTC 60

QY 61 GAGCTGTGGAACCTGAAGAGTTCAAGTTATACCTGGGACCGGACAGAGACTGGAGAA 120
DB 61 GAGCTGTGGAACCTGAAGAGTTCAAGTTATACCTGGGACCGGACAGAGACTGGAGAA 120

QY 121 GGCAAGATCCCTGGGAGAGCATGGAAGCCCGGTCCCTGGGAAATGGCCAGCTGCTC 180
DB 121 GGCAAGATCCCTGGGAGAGCATGGAAGCCCGGTCCCTGGGAAATGGCCAGCTGCTC 180

QY 181 ATACCCACCTTCGGGCGCAGAGAGGCGCTGGAGGTTGCTCTCAGACCTTTGAGGGATA 240
DB 181 ATACCCACCTTCGGGCGCAGAGAGGCGCTGGAGGTTGCTCTCAGACCTTTGAGGGATA 240

QY 241 AACAGGAAGGACTGTGGGAGAGAGGACAGAGAGGACCTGGTGAGG- 288
DB 241 AACAGGAAGGACTGTGGGAGAGAGGACAGAGAGGACCTGGTGAGG- 288

QY 289 - 288
DB 289 - 288

QY 301 GGTGGCCGCTCTCACTTGGGAACAGTCAACATGCTTCTGGAACTCTCTTGTCACT 360
DB 301 GGTGGCCGCTCTCACTTGGGAACAGTCAACATGCTTCTGGAACTCTCTTGTCACT 360

QY 289 -GATCCCGCAGGAACTTACAGGAGCTATGTCCGAGGAAATCCCGCTCATG 339
DB 361 CCAAGAAAGATCCCGCAGGAACTTACAGGAGCTATGTCCGAGGAAATCCCGCTCATG 420

QY 340 GAAGACCGCAATCGCCCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTC 399
DB 421 GAAGACCGCAATCGCCCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGGCTC 480

QY 400 CTGCTGTGAGGAGGACTCAACCCCATGAGAGTCCAGAGGAGGAGTCTTGGACACAGGC 459
DB 481 CTGCTGTGAGGAGGACTCAACCCCATGAGAGTCCAGAGGAGGAGTCTTGGACACAGGC 540

QY 460 CGGGGACAGCGGAGCGCTGGGACACAGGCTAGCCCTCATCAAGATAGAGACCTCTTT 519
DB 541 CGGGGACAGCGGAGCGCTGGGACACAGGCTAGCCCTCATCAAGATAGAGACCTCTTT 600

QY 520 GAGCCAGACGAGGAGCGCCCGGAGCCACCGGCAACCGTGTGATGCAAGGCGCGCAGGG 579
DB 601 GAGCCAGACGAGGAGCGCCCGGAGCCACCGGCAACCGTGTGATGCAAGGCGCGCAGGG 660

QY 580 ATAGGCAAGTCCATGCTGGGACACAGAGTATGCTGAGCTGGGCGGAGGAGCTCTTC 639
DB 661 ATAGGCAAGTCCATGCTGGGACACAGAGTATGCTGAGCTGGGCGGAGGAGCTCTTC 720

QY 640 CAAGGCAAGTTCATATCTCTTACATCAACTGCAAGGAGGATGAACAGAGTGCCACG 699
DB 721 CAAGGCAAGTTCATATCTCTTACATCAACTGCAAGGAGGATGAACAGAGTGCCACG 780

QY 700 GAATGCAAGTCAAGACCTCATCTTCACTGCTGCTGCTGCTGGGAGGAGAAACGCCCAAG 759
DB 781 GAATGCAAGTCAAGACCTCATCTTCACTGCTGCTGCTGCTGGGAGGAGAAACGCCCAAG 840

QY 760 GAGCTCATCCGAGTTCGCGAGCGCTCTTTTCAATCATGACCGCTTCGATGAGCTCAAG 819
DB 841 GAGCTCATCCGAGTTCGCGAGCGCTCTTTTCAATCATGACCGCTTCGATGAGCTCAAG 900

QY 820 CCTTCTTTTCCAGCATCTCAGGAGACCTGTGTGCTCTGCTGGGAGGAGAAACGCCCAAG 879
DB 901 CCTTCTTTTCCAGCATCTCAGGAGACCTGTGTGCTCTGCTGGGAGGAGAAACGCCCAAG 960

QY 880 GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATC 939
DB 961 GAGCTGCTTCTTAACAGCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATC 1020

QY 940 ACCACAGCGGCCACCGGCTTTGGAGAGCTCCACGCTGTGTGGAGCAACCCAGGCACTGTG 999
DB 1021 ACCACAGCGGCCACCGGCTTTGGAGAGCTCCACGCTGTGTGGAGCAACCCAGGCACTGTG 1080

QY 1000 GAGATCCTGGGCTTCTCTGAGGCAAGAGGAGAAATATCTTCAAGTATTTCCACAAT 1059
DB 1081 GAGATCCTGGGCTTCTCTGAGGCAAGAGGAGAAATATCTTCAAGTATTTCCACAAT 1140

QY 1060 GCAGAGAGGCGGCGCCCAAGTCTTCAAATTACGTGAGGACAAACGAGCCTCTCTTCCACCATG 1119
DB 1141 GCAGAGAGGCGGCGCCCAAGTCTTCAAATTACGTGAGGACAAACGAGCCTCTCTTCCACCATG 1200

QY 1120 TGCTTGTCCCTCTGGTGTGCTGGGTGTGTACTGCTCTCCAGCAGCAGCTGAGGGT 1179
DB 1201 TGCTTGTCCCTCTGGTGTGCTGGGTGTGTACTGCTCTCCAGCAGCAGCTGAGGGT 1260

QY 1180 GGGGGCTGTGTAGACAGAGCTCCAGGACCAACCACTGCTGCTGCTGCTGCTGCTGCTG 1239
DB 1261 GGGGGCTGTGTAGACAGAGCTCCAGGACCAACCACTGCTGCTGCTGCTGCTGCTGCTG 1320

QY 1240 AGTCTGATGCAACCCCAAGCGGCGGCCCTCCAGCCTCCAGCCTCCCAACCAACGAGAGGG 1299
DB 1321 AGTCTGATGCAACCCCAAGCGGCGGCCCTCCAGCCTCCAGCCTCCCAACCAACGAGAGGG 1380

QY 1300 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1359
DB 1381 TTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGAC 1440

QY 1360 CTCGGAGACAGCGCTAGACGGGAGAGCTCTGCTGCTCTCTCAACATGAATCTTC 1419
DB 1441 CTCGGAGACAGCGCTAGACGGGAGAGCTCTGCTGCTCTCTCAACATGAATCTTC 1500

QY 1420 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTATCCACTTGTGAGTTTCCAGGAATTC 1479
DB 1501 CAGAAGGACATCAACTGTGAGAGTACTACAGCTTCTATCCACTTGTGAGTTTCCAGGAATTC 1560

QY 1480 TTTGACGCTATGTACTATATCTTGGACGAGGGGAGGGGGGAGGAGGAGGAGGAGGAGG 1539
DB 1561 TTTGACGCTATGTACTATATCTTGGACGAGGGGAGGGGGGAGGAGGAGGAGGAGGAGG 1620

QY 1540 GTGACAGGCTGTGTGACGAGTACGCTTTTCTGAAAGGAGCTTCTGAGGACTTCCAGC 1599
DB 1621 GTGACAGGCTGTGTGACGAGTACGCTTTTCTGAAAGGAGCTTCTGAGGACTTCCAGC 1680

QY 1600 CGCTTCTCTTGTGAGCTCTTGAACAGGAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1659
DB 1681 CGCTTCTCTTGTGAGCTCTTGAACAGGAGACCAAGAGGAGGAGGAGGAGGAGGAGGAGG 1740

QY 1660 TGAAGGCTCTGCGCGACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAGAGCTCAG 1719
DB 1741 TGAAGGCTCTGCGCGACATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAGAGCTCAG 1800

QY 1720 AGCGAGGCTCCACCTGACAGGAGGCTCTGAGGTTCTTCACTGCTGCTGCTGCTGCTGCTG 1779
DB 1801 AGCGAGGCTCCACCTGACAGGAGGCTCTGAGGTTCTTCACTGCTGCTGCTGCTGCTGCTG 1860

QY 1780 CAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCTTCCAGGAGTGTGCTGCTGCTGCTGCTG 1839
DB 1861 CAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCCTTCCAGGAGTGTGCTGCTGCTGCTGCTG 1920

QY 1840 ATTGCTCTCAAGATGAGGACATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1899
DB 1921 ATTGCTCTCAAGATGAGGACATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980

QY 1900 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1959
DB 1981 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2040

QY 1960 TCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2019
DB 2041 TCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2097

QY 2020 GCCTACAGTGAACATCTGGGAGGCGGCGCTGTGCAACCACTCCAAACCTGTAGAGCTGTCT 2079

Db 2098 GCTACAGTGAACATCTGGCAGCGGCTGTGCAACCAATCCAAACCTGATAGAGTGTCT 2157
|
QY 2080 CTGTACGGAATGCCCTGGCGAGCCGGGGGTGAAGTCTGTCTGAAGACTCAGACAC 2139
|
Db 2158 CTGTACGGAATGCCCTGGCGAGCCGGGGGTGAAGTCTGTCTGAAGACTCAGACAC 2217
|
QY 2140 CCCAATGCAAACTTCAGAACTTGAGGCTGAGAGGTGCGGATCTCCAGCTCAGCTGC 2199
|
Db 2218 CCCAATGCAAACTTCAGAACTTGAGGCTGAGAGGTGCGGATCTCCAGCTCAGCTGC 2277
|
QY 2200 GAGGACCTCTGCGAGCTCTATAGCAATAAGAAATTTGCAAGAGTGAATCTCAGTGC 2259
|
Db 2278 GAGGACCTCTGCGAGCTCTATAGCAATAAGAAATTTGCAAGAGTGAATCTCAGTGC 2337
|
QY 2260 AACGGGTTGGATCCAGGATGATGCTTGGGAGGCTGGGATCCAGTGC 2319
|
Db 2338 AACGGGTTGGATCCAGGATGATGCTTGGGAGGCTGGGATCCAGTGC 2397
|
QY 2320 AGGCTGAGATGATTCAGTTGAGGAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGTG 2379
|
Db 2398 AGGCTGAGATGATTCAGTTGAGGAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGTG 2457
|
QY 2380 GCTTCTGTGCTGGCACCACCACTGTTGAGTTCGAGCTGACAGGAAATGCACTG 2439
|
Db 2458 GCTTCTGTGCTGGCACCACCACTGTTGAGTTCGAGCTGACAGGAAATGCACTG 2517
|
QY 2440 GAGGATTTGGGCTGAGGTTACTATCCAGGAGTCCAGGATCCAGTCTGCACTACG 2499
|
Db 2518 GAGGATTTGGGCTGAGGTTACTATCCAGGAGTCCAGGATCCAGTCTGCACTACG 2577
|
QY 2500 ACTTGTGTGCTGAAGATCTGCGGCTCACTGTGCTGCTGACAGCTGGCTCACT 2559
|
Db 2578 ACTTGTGTGCTGAAGATCTGCGGCTCACTGTGCTGCTGACAGCTGGCTCACT 2637
|
QY 2560 CTCAGTGTGAACAGAGCTGAGAGCTGAGCTGAGCTGAATGAGTGGGGACCTC 2619
|
Db 2638 CTCAGTGTGAACAGAGCTGAGAGCTGAGCTGAGCTGAATGAGTGGGGACCTC 2697
|
QY 2620 GGGTGTGCTGCTGTGTGAGGCTCAGGATCCAGTCCAGTCCAGTCCAGTCCG 2679
|
Db 2698 GGGTGTGCTGCTGTGTGAGGCTCAGGATCCAGTCCAGTCCAGTCCAGTCCG 2757
|
QY 2680 TTGGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2739
|
Db 2758 TTGGGCTGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 2817
|
QY 2740 AACCAACCTCCGGAGCTGAGTGTGAGTGTCAACGACCTGGGAGCTGGGCTGTGG 2799
|
Db 2818 AACCAACCTCCGGAGCTGAGTGTGAGTGTCAACGACCTGGGAGCTGGGCTGTGG 2877
|
QY 2800 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGACTCCAGAACTGGCTGATAGC 2859
|
Db 2878 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGACTCCAGAACTGGCTGATAGC 2937
|
QY 2860 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGACTCCAGAACTGGCTGATAGC 2919
|
Db 2938 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGACTCCAGAACTGGCTGATAGC 2997
|
QY 2920 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGGAGTCCAGTGTCCAGTGTCTTGC 2979
|
Db 2998 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGGAGTCCAGTGTCCAGTGTCTTGC 3057
|
QY 2980 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGGAGTCCAGTGTCCAGTGTCCAGTGT 3039
|
Db 3058 TTGCTGGCTGAGGCTGCAACCTCCGCTGAGGAGTCCAGTGTCCAGTGTCCAGTGT 3117
|
QY 3040 AATAAATGACCAAGTGTGGAGGCTTCCAGTGTCCAGTGTCCAGTGTCCAGTGT 3099
|
Db 3118 AATAAATGACCAAGTGTGGAGGCTTCCAGTGTCCAGTGTCCAGTGTCCAGTGT 3177
|
QY 3100 GCTGCTGA 3108
|

3178 GCTGCTGA 3186

RESULT 4
ADC30316

ID ADC30316 standard; cDNA; 3306 BP.

XX AC ADC30316;

XX DT 18-DEC-2003 (first entry)

XX Human novel cDNA sequence, SEQ ID NO:398.

DE Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; anti-anaemic; anticoagulant; thrombolytic; vulnary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 19; gene; ss.

XX Homo sapiens.

XX WO2003029271-A2.

XX 10-APR-2003.

XX 24-SEP-2002; 2002WO-US030474.

XX 24-SEP-2001; 2001US-0324631P.

XX (HYSE-) HYSEQ INC.

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;
PI Haley-Vicente D, Drmanac RT;

DR WPI; 2003-371981/35.

DR P-PSDB; ADC31287.

XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.

PS Claim 1; SEQ ID NO 398; 1185bp; English.

XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or

CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC fp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3306 BP; 718 A; 958 C; 939 G; 691 T; 0 U; 0 Other;
Query Match 78.8%; Score 2449.4; DB 9; Length 3306;
Best Local Similarity 93.8%; Pred. No. 0;
Matches 2644; Conservative 0; Mismatches 1; Indels 174; Gaps 2;
290 ATCCCCAGGAACCTACAGGACTATGTCCAGAGAAATTCGGCTCATGGAAGACCGCA 349
Db ATCCCCAGGAACCTACAGGACTATGTCCAGAGAAATTCGGCTCATGGAAGACCGCA 557
350 ATGCGCGCTAGGGAATGTCTAACTCAGCACCCTGATACCCGGCTCTGCTGTGA 409
Db ATGCGCGCTAGGGAATGTCTAACTCAGCACCCTGATACCCGGCTCTGCTGTGA 617
410 AGGAGCACTCAAAACCCCATGCAAGTCCAGAGAGCTTCTGGACACAGGCCGGGACACG 469
Db AGGAGCACTCAAAACCCCATGCAAGTCCAGAGAGCTTCTGGACACAGGCCGGGACACG 677
470 CGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGCG 529
Db CGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGCG 737
530 AGGAGCCCGGAGCCAGCCGACCTGTGACTGGGCGGACGGAGCTCTTCAAGCGAGAT 589
Db AGGAGCCCGGAGCCAGCCGACCTGTGACTGGGCGGACGGAGCTCTTCAAGCGAGAT 797
590 CCATGCTGGGACACAAAGTGATGCTGACTGGGCGGACGGAGCTCTTCAAGCGAGAT 649
Db CCATGCTGGGACACAAAGTGATGCTGACTGGGCGGACGGAGCTCTTCAAGCGAGAT 857
650 TTGATTAATCTTCTCATCAATTCAGGGAGATGAACAGAGTGCCACGGAATGCAGCA 709
Db TTGATTAATCTTCTCATCAATTCAGGGAGATGAACAGAGTGCCACGGAATGCAGCA 917
710 TGCAGACCTCATCTTTCAGTGTGGCTGAGCCAGCGGCTCTCCAGAGCTCATCC 769
Db TGCAGACCTCATCTTTCAGTGTGGCTGAGCCAGCGGCTCTCCAGAGCTCATCC 977
770 GAGTTCGCGAGCGCTCTCTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCC 829
Db GAGTTCGCGAGCGCTCTCTTTCATCATCGACGGCTTCGATGAGCTCAAGCTTCTTTCC 1037
830 ACGATCTCAGGACCTGTGCTCTGCTGGGAGGAGAAACGCGCCACGAGTGCTTC 889
Db ACGATCTCAGGACCTGTGCTCTGCTGGGAGGAGAAACGCGCCACGAGTGCTTC 1097
890 TTAACAGCTTAATTCGGAAGAGCTCTCCCTGAGTATCTTCTCATCACACAGCG 949
Db TTAACAGCTTAATTCGGAAGAGCTCTCCCTGAGTATCTTCTCATCACACAGCG 1157
950 CCAGCGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCAGGCAATGGAGATCTGG 1009
Db CCAGCGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCAGGCAATGGAGATCTGG 1217
1010 GCTTCTCTGAGGAGAGAGAAATCTTCTCAAGATTTTCAAGTTCAGAGCAGG 1069
Db GCTTCTCTGAGGAGAGAGAAATCTTCTCAAGATTTTCAAGTTCAGAGCAGG 1277
1070 CGGCGCAAGTCTCAATTCAGTGGAGGACACAGGCTCTTCCACATGCTTCGCTCC 1129
Db CGGCGCAAGTCTCAATTCAGTGGAGGACACAGGCTCTTCCACATGCTTCGCTCC 1337
1129 CGGCGCAAGTCTCAATTCAGTGGAGGACACAGGCTCTTCCACATGCTTCGCTCC 1189
Db CGGCGCAAGTCTCAATTCAGTGGAGGACACAGGCTCTTCCACATGCTTCGCTCC 1397
1130 CCTGTGTGTGTGGT 1189
Db CCTGTGTGTGTGGT 1397

1190 TGAGACAGAGCTTCAGGACCAACCACTCAGTGTACATGCTTACTCTGAGTCTGATGC 1249
Db TGAGACAGAGCTTCAGGACCAACCACTCAGTGTACATGCTTACTCTGAGTCTGATGC 1457
1250 AACCCAAAGCCGGGGCCCGGCTCCAGCCCCCAGCCCCAACCCAGAGAGGTTGCTCCT 1309
Db AACCCAAAGCCGGGGCCCGGCTCCAGCCCCCAGCCCCAACCCAGAGAGGTTGCTCCT 1517
1310 TGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTAGGAGCAGACCTCCGGAAGC 1369
Db TGGCGGAGATGGGCTCTGGAATCAGAAATCTTATTAGGAGCAGACCTCCGGAAGC 1577
1370 ACGGCTTAGAGCGGGAAGAAGTCTCTGCTCTCTCAATGAAATCTTCCAGAGGACA 1429
Db ACGGCTTAGAGCGGGAAGAAGTCTCTGCTCTCTCAATGAAATCTTCCAGAGGACA 1637
1430 TCAACTGTGAGAGTACTACAGCTTCACTGAGTGTTCAGGAATCTTTCGACGTA 1489
Db TCAACTGTGAGAGTACTACAGCTTCACTGAGTGTTCAGGAATCTTTCGACGTA 1697
1490 TGTACTATATCTCGACGAGGGGAGGGCGGGGAGGCCAGAGCCAGAGCCTGACCAAGC 1549
Db TGTACTATATCTCGACGAGGGGAGGGCGGGGAGGCCAGAGCCAGAGCCTGACCAAGC 1757
1550 TGTGACCGAGTACGCGTCTTCTGAAAGAGCTTCTGGAAGTCTTCCAGCGCTTCTGT 1609
Db TGTGACCGAGTACGCGTCTTCTGAAAGAGCTTCTTCCAGCGCTTCTGTGGAAGTCT 1817
1610 TTGGAGCTCTGAGAGGAGCAGGAGCCAGGAGCCCTGAGAGAGTCTCTGCTGAGAGTCT 1669
Db TTGGAGCTCTGAGAGGAGCAGGAGCCAGGAGCCCTGAGAGAGTCTCTGCTGAGAGTCT 1877
1670 CGCCGCAATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAAAGCTCAGAGCCAGCCT 1729
Db CGCCGCAATCAAGATGGAACCTGTTGAGTGGATCCAAAGCAAAAGCTCAGAGCCAGCCT 1937
1730 CCACCTTCGACAGAGGCTCTTGGAGTCTTTCAGTGTCTGACGATCCAGATCCAGAGGAG 1789
Db CCACCTTCGACAGAGGCTCTTGGAGTCTTTCAGTGTCTGACGATCCAGATCCAGAGGAG 1997
1790 AGTTTATCCAGAGCCCTGAGCCACTTCCAGGATGCTGCTGAGAGAGTCTGCTGAGAGTCT 1849
Db AGTTTATCCAGAGCCCTGAGCCACTTCCAGGATGCTGCTGAGAGAGTCTGCTGAGAGTCT 2057
1850 AGATGAGCAGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1909
Db AGATGAGCAGATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2117
1910 ACTTGTATGGGCGCACTTACAGCCGGAAGAGACCGCGGAGGCTGCTGCTGCTGCTGCTGCT 1969
Db ACTTGTATGGGCGCACTTACAGCCGGAAGAGACCGCGGAGGCTGCTGCTGCTGCTGCTGCT 2177
1970 CGCACAGCTGTTGGTGCAGCTCAGACAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2029
Db CGCACAGCTGTTGGTGCAGCTCAGACAGAGAGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2234
2030 AACATCTGGAGCGGCTCTGCAACCAATCAAACTGATAGAGCTGCTCTGCTGCTGCTGCTGCT 2089
Db AACATCTGGAGCGGCTCTGCAACCAATCAAACTGATAGAGCTGCTCTGCTGCTGCTGCTGCT 2294
2090 ATGCTTGGGAGCGCGGGGGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2149
Db ATGCTTGGGAGCGCGGGGGTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2354
2150 AACTTCAAGAACCTGAGGCTGAAGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2209
Db AACTTCAAGAACCTGAGGCTGAAGAGTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2414
2210 CTGAGCTCTCATAGCCAAATGAAGATTTGAAGAGTGGATCTCAGTGGCAACCGGCTG 2269
Db CTGAGCTCTCATAGCCAAATGAAGATTTGAAGAGTGGATCTCAGTGGCAACCGGCTG 2474

atopic dermatitis; arthritis; bacterial infection; viral infection;
parasitic infection; protozoal infection; fungal infection.

Homo sapiens.

WO200272630-A2.

19-SEP-2002.

07-FEB-2002; 2002WO-US003844.

09-FEB-2001; 2001US-0268118P.

21-FEB-2001; 2001US-0270963P.

22-FEB-2001; 2001US-0270858P.

23-FEB-2001; 2001US-0271194P.

07-MAR-2001; 2001US-0274071P.

12-APR-2001; 2001US-0283496P.

09-NOV-2001; 2001US-0344650P.

(INCY-) INCYTE GENOMICS INC.

Thornnton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;

Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan PA;

Gandhi AR, Ding L, Yue H, Gietzen KJ, Walia NK, Thangavelu K;

Elliott VS, Marquis JP;

WPI; 2002-723320/78.

P-PSDB; ABG97475.

New human nucleic acid-associated proteins (NAAP), useful for diagnosing, treating and preventing diseases or conditions associated with the aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis, infections.

Claim 5; Page 160; 162pp; English.

The invention relates to an isolated polypeptide comprising one of 10 human nucleic acid associated protein (NAAP1-10), or a biologically active or immunogenic fragment of the polypeptide, and their encoding nucleic acid. Also included are a recombinant polynucleotide comprising a promoter sequence operably linked to the polynucleotide, a cell transformed with the recombinant polynucleotide, a transgenic organism comprising the recombinant polynucleotide, an anti-NAAP antibody, screening for a compound that is effective as an ant/agonist or modulator of NAAP, generating an expression profile of a sample containing the polynucleotides and an array comprising different nucleotide molecules affixed on a solid substrate, nucleotide molecule comprises a first oligonucleotide or polynucleotide sequence specifically hybridizable with at least 30 contiguous nucleotides of the target (NAAP) polynucleotide. The polypeptides and polynucleotides are useful in diagnosing, treating and preventing diseases or conditions associated with the decreased expression or overexpression of NAAP, such as cell proliferative diseases (e.g. cancer, atherosclerosis, hepatitis), neurological disorders (e.g. Parkinson's disease, Alzheimer's disease, stroke, epilepsy), developmental disorders (renal tubular acidosis, anaemia, glaucoma, hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many other diseases and disorders listed in the specification. These are also useful in assessing the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of NAAP. The NAAP or its fragments are useful in screening compounds for effectiveness as agonist or antagonist of the polypeptides, or in altering the expression of the target polynucleotide and compounds that specifically bind to or modulate the activity of the polypeptide. The microarray is useful in monitoring or measuring protein-protein interactions, drug-target interactions, and gene expression profiles. The present sequence encodes an NAAP protein

Sequence 2158 BP; 489 A; 578 C; 639 G; 452 T; 0 U; 0 Other;

Query Match 54.1%; Score 1691.8; DB 6; Length 2158;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;

QY 1408 ATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTTCATCCACTTGACT 1467
DB 1 ATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTTCATCCACTTGACT 60
QY 1468 TTCCAGGAAATCTTTGCGAGCTATGTAATATATCTCCGACGAGGGGAGGGCGGGCAGGC 1527
DB 61 TTCCAGGAAATCTTTGCGAGCTATGTAATATATCTCCGACGAGGGGAGGGCGGGCAGGC 120
QY 1528 CCAGACCCAGGAGCTGACACGAGGCTGTGACCGAGTACGGCTTTTCTGAAGAGAGCTTCCTG 1587
DB 121 CCAGACCCAGGAGCTGACACGAGGCTGTGACCGAGTACGGCTTTTCTGAAGAGAGCTTCCTG 180
QY 1588 GCACATCACAGCGCTTCTGTTTCGACTCTCTGTAACGAGGAGACAGGAGCCACCTGGAG 1647
DB 181 GCACATCACAGCGCTTCTGTTTCGACTCTCTGTAACGAGGAGACAGGAGCCACCTGGAG 240
QY 1648 AAGAGTCTCTGCTGGAAGTCTTCGCCGACATCAAGATGGACCTGTTCAGTGGATCCAA 1707
DB 241 AAGAGTCTCTGCTGGAAGTCTTCGCCGACATCAAGATGGACCTGTTCAGTGGATCCAA 300
QY 1708 AGCAAGCTCAGAGGAGCGCTCCACCTCGACGAGGGCTCCTTGGAGTTCCTTCAGCTGC 1767
DB 301 AGCAAGCTCAGAGGAGCGCTCCACCTCGACGAGGGCTCCTTGGAGTTCCTTCAGCTGC 360
QY 1768 TTGTACGAGATCCAGGAGGAGGTTTATCCAGCAGGCGCTGAGCCACTTCCAGGTGATC 1827
DB 361 TTGTACGAGATCCAGGAGGAGGTTTATCCAGCAGGCGCTGAGCCACTTCCAGGTGATC 420
QY 1828 GTGGTCAGCAACATTCCTCCAGATGGAGCACATGGTCTCCTCGTTCTGTCTGAAGCGC 1887
DB 421 GTGGTCAGCAACATTCCTCCAGATGGAGCACATGGTCTCCTCGTTCTGTCTGAAGCGC 480
QY 1888 TCCAGGAGCGCCAGGCTGCTGCACTTGTATGGCGCACCTACAGCGCGGACCGGGAAGAC 1947
DB 481 TCCAGGAGCGCCAGGCTGCTGCACTTGTATGGCGCACCTACAGCGCGGACCGGGAAGAC 540
QY 1948 CGCGGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACACAGAGGAGCC 2007
DB 541 CGCGGAGGTGCTCCGAGGAGCGCACAGCTGTTGGTGCAGCT---ACGAGAGGAGCC 597
QY 2008 GTTCTGCTGGAGCGCTACAGTGAAACATCTGGCAGCGCGCCTGTGCACCAATCAAACTG 2067
DB 598 GTTCTGCTGGAGCGCTACAGTGAAACATCTGGCAGCGCGCCTGTGCACCAATCAAACTG 657
QY 2068 ATAGAGCTGCTCTGTACGAAATCCCTTGGCAGCGCGGCGGTGAAGTCTCTGTCAA 2127
DB 658 ATAGAGCTGCTCTGTACGAAATCCCTTGGCAGCGCGGCGGTGAAGTCTCTGTCAA 717
QY 2128 GGACTTCAGACACCCCACTGCAAACTTCAGAACCTGAGGCTCAAGAGGTCCGCATCTCC 2187
DB 718 GGACTTCAGACACCCCACTGCAAACTTCAGAACCTGAGGCTCAAGAGGTCCGCATCTCC 777
QY 2188 AGCTCAGGCTGCGAGGAGCTCTCTGCAGCTCTCATAGCAATAAGAAATTCAGAGGATG 2247
DB 778 AGCTCAGGCTGCGAGGAGCTCTCTGCAGCTCTCATAGCAATAAGAAATTCAGAGGATG 837
QY 2248 GATCTCAGTGGCAACGGGCTTGAGATCCAGGAGTATGCTCTTTCGAGGCGCTGCGG 2307
DB 838 GATCTCAGTGGCAACGGGCTTGAGATCCAGGAGTATGCTCTTTCGAGGCGCTGCGG 897
QY 2308 CATCCCCAGTGCAGGCTCAGATGATTCAGTGGAGAGTGTCACTGGAGTCCGGGCT 2367
DB 898 CATCCCCAGTGCAGGCTCAGATGATTCAGTGGAGAGTGTCACTGGAGTCCGGGCT 957
QY 2368 TGTCAAGAGATGGCTTCTGTGCTCGGACCAACCCACATCTGGTGGATGGACCTGACA 2427
DB 958 TGTCAAGAGATGGCTTCTGTGCTCGGACCAACCCACATCTGGTGGATGGACCTGACA 1017
QY 2428 GGAAATGACCTGGAGGATTTGGGCTGAGGTATCTATGCCAGGAGTCCAGGACCCAGTGC 2487
DB 1018 GGAAATGACCTGGAGGATTTGGGCTGAGGTATCTATGCCAGGAGTCCAGGACCCAGTGC 1077

QY 649 TTGATTATCTCTTACATCACTGCAGGGAGATGAACACAGATGCCACGGAATGCAGC 708
 Db 1314 -----AGGGAGATGAACACAGATGCCACGGAATGCAGC 1346
 QY 709 ATGCAAGACCTCATCTTTCAGCTGCTGGCCCTGAGCCAGCGGCCCTCTCCAGAGCTCATC 768
 Db 1347 ATGCAAGACCTCATCTTTCAGCTGCTGGCCCTGAGCCAGCGGCCCTCTCCAGAGCTCATC 1406
 QY 769 CGAGTTCCGAGCGCTCTCTTTTCATCATCGACGGCTTCGATGAGTCAAGCTTCTTTTC 828
 Db 1407 CGAGTTCCGAGCGCTCTCTTTTCATCATCGACGGCTTCGATGAGTCAAGCTTCTTTTC 1466
 QY 829 CACGATCTCAGGACCTCTGTCCTCTGCTGGAGGAGAAACGGCCACCGAGCTGCTT 888
 Db 1467 CACGATCTCAGGACCTCTGTCCTCTGCTGGAGGAGAAACGGCCACCGAGCTGCTT 1526
 QY 889 CTTAAGACCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACAGG 948
 Db 1527 CTTAAGACCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACAGG 1586
 QY 949 CCCACGGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGACATCTGGAGATCCTG 1008
 Db 1587 CCCACGGCTTTGGAGAAGCTCCACCGTCTGCTGGAGCACCCAGGACATCTGGAGATCCTG 1646
 QY 1009 GCTTCTCTGAGGAGAAAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1068
 Db 1647 GCTTCTCTGAGGAGAAAGGAATACTTCTACAAGTATTTCCACAATGCAGAGCAG 1706
 QY 1069 GGGGCCAAGTCTCAATTAAGTAGGAGAACAGAGCTCTCTTCAACATGCTTCTGTC 1128
 Db 1707 GGGGCCAAGTCTCAATTAAGTAGGAGAACAGAGCTCTCTTCAACATGCTTCTGTC 1766
 QY 1129 CCCCTGGTGTGCTGGTGTGATGCTGCTGCTGAGGAGCACCTGAGGCTGGGGGGCTG 1188
 Db 1767 CCCCTGGTGTGCTGGTGTGATGCTGCTGCTGAGGAGCACCTGAGGCTGGGGGGCTG 1826
 QY 1189 TTGACAGACGCTCAGGACCACTGAGTGTACATGCTCTACCTGCTGAGTCTGATG 1248
 Db 1827 TTGACAGACGCTCAGGACCACTGAGTGTACATGCTCTACCTGCTGAGTCTGATG 1886
 QY 1249 CAACCCAGCCGGGGCCCGCGCTCCAGCCGCCCAACCCAGAGAGGTTGTGCTCC 1308
 Db 1887 CAACCCAGCCGGGGCCCGCGCTCCAGCCGCCCAACCCAGAGAGGTTGTGCTCC 1946
 QY 1309 TTGGGGCAGATGGCTCTGGAATCAGAAAATCTTATTTGAGGAGCAGACCTCCGGAG 1368
 Db 1947 TTGGGGCAGATGGCTCTGGAATCAGAAAATCTTATTTGAGGAGCAGACCTCCGGAG 2006
 QY 1369 CACGCTCTAGACGGGGAAGAGCTCTGCTCTTCTCAACATGAACATCTTCCAGAGGAC 1428
 Db 2007 CACGCTCTAGACGGGGAAGAGCTCTGCTCTTCTCAACATGAACATCTTCCAGAGGAC 2066
 QY 1429 ATCACTGTGAGGATCTACAGCTTCTCACTGAGTTTCCAGGAATTTCTTTCAGCT 1488
 Db 2067 ATCACTGTGAG----- 2078
 QY 1489 ATGTACTATATCTTGGACGAGGGGAGGGCGGGCAGGCCACAGACGAGACGACGAG 1548
 Db 2079 ----- 2078
 QY 1549 CTGTTGACCGAGTACGCGTTTCTGAAAGAGCTTCTTGGCACTCACCAGCGCTTCTG 1608
 Db 2079 -----AGGAGCTTCTTGGCACTCACCAGCGCTTCTG 2111
 QY 1609 TTGGAATCTTGAACGAGGAGACGAGGAGCCACTGGAGAGAGTCTCTGCTGGAAGTTC 1668
 Db 2112 TTGGAATCTTGAACGAGGAGACGAGGAGCCACTGGAGAGAGTCTCTGCTGGAAGTTC 2171
 QY 1669 TCGCGCACATCAAGTACGCTGTCAGTGGATCCAAAGCAAGAGCTCAGAGCGACGGC 1728
 Db 2172 TCGCGCACATCAAGTACGCTGTCAGTGGATCCAAAGCAAGAGCTCAGAGCGACGGC 2231
 QY 1729 TCCACCTGACAGCGGCTCTTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG 1788

Db 2232 TCCACCTGACAGCGCTCTTTCAGTCTTTCAGTCTGTTACGAGATCCAGGAGGAG 2291
 QY 1789 GAGTTTATCCAGCAGCGCTTTCAGTCTTTCAGTCTGTTACGAGATCCAGGAGGAG 1848
 Db 2292 GAGTTTATCCAGCAGCGCTTTCAGTCTTTCAGTCTGTTACGAGATCCAGGAGGAG 2351
 QY 1849 AAGATGGAGCACATGCTCTCTGCTTTCAGTCTGTTACGAGATCCAGGAGGAG 1908
 Db 2352 AAGATGGAGCACATGCTCTCTGCTTTCAGTCTGTTACGAGATCCAGGAGGAG 2411
 QY 1909 CACTTGTATGGCCACCTTACAGCGGAGCGGGGAGAGCCGCGCAGAGTGTCTCCGAGGA 1968
 Db 2412 CACTTGTATGGCCACCTTACAGCGGAGCGGGGAGAGCCGCGCAGAGTGTCTCCGAGGA 2471
 QY 1969 GGGCAGCAGCTGTTGGTTCAGCTCAGACGAGAGGAGCGGTTCTGCTGAGACGCTTACAGT 2028
 Db 2472 GGGCAGCAGCTGTTGGTTCAGCT---ACCAGAGAGGAGCGTTCTGCTGAGACGCTTACAGT 2528
 QY 2029 GAAATCTGGCAGCGGCTGTCACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA 2088
 Db 2529 GAAATCTGGCAGCGGCTGTCACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA 2588
 QY 2089 AATGCCCTGGCAGCGGCGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCAACTGC 2148
 Db 2589 AATGCCCTGGCAGCGGCGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCAACTGC 2648
 QY 2149 AATCTTCAAGACTCA----- 2164
 Db 2649 AATCTTCAAGACTCAAGGTTAAATTTATCATATATAACATGATATTTGAAATAAATATA 2708
 QY 2165 ----- 2164
 Db 2709 TTGGCCAGTATGATGGCTCAGCGCTGTAATTCAGACACTTTGGGAGGCCAGATGGGA 2768
 QY 2165 ----- 2164
 Db 2769 GGAATCTTGCAGGAGTTCAAGCAGCGCTGCGCAACATGGTGAACCCCACTCTCTA 2828
 QY 2165 ----- 2164
 Db 2829 CTAAATAACCAAAATGAGCCAGGCTGCTGGCAGACAGCTCTGTAAGCCAGCTACTCAGG 2888
 QY 2165 -----GGCTGAAGAGTGCC 2179
 Db 2889 AGGCCAAGCAGGAGGATGCTTCAACCCAGGAGCAGAGGTTGCTGCTGAAGAGTGCC 2948
 QY 2180 GCATCTCCAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCCCAATAAGAAATTTGA 2239
 Db 2949 GCATCTCCAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCCCAATAAGAAATTTGA 3008
 QY 2240 CAGGATGGATCTCAGTGGCAACCGCGTGGATTTCCAGGCGATGATGCTTTCGAGG 2299
 Db 3009 CAGGATGGATCTCAGTGGCAACCGCGTGGATTTCCAGGCGATGATGCTTTCGAGG 3068
 QY 2300 GCCTGCGCATCCCGAGTGCAGGCTGAGAGTATTCAGTTCAGGAGAGTCTCAGCTGGAGT 2359
 Db 3069 GCCTGCGCATCCCGATGCGAGCTGAGAGTATTCAGTTCAGGAGAGTCTCAGCTGGAGT 3128
 QY 2360 CCGGGCTTGTGAGGATGGCTTCTGCTCGGCACCAACCCACATCTGTTGAGTTGG 2419
 Db 3129 CCGGGCTTGTGAGGATGGCTTCTGCTCGGCACCAACCCACATCTGTTGAGTTGG 3188
 QY 2420 ACCTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCGAGGAGCTGAGGC 2479
 Db 3189 ACCTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCGAGGAGCTGAGGC 3248
 QY 2480 ACCAGTCTGAGACTACGAGCTTGTGCTGAGATCTGCGCTCACTGCTGCTGCT 2539
 Db 3249 ACCAGTCTGAGACTACGAGCTTGTGCTGAGATCTGCGCTCACTGCTGCTGCT 3308
 QY 2540 GTGACGAGCTGCGCTCAACTCTCAGTGTGAACCCAGAGCTGAGAGGCTGAGGAGCTGAGGC 2599

Db 3309 GTGACGAGCTGGCTCACTCTCAAGTGTGAACACGAGCCCTGAGAGAGCTGACCTGAGCC 3368
 QY 2600 TGAATGAGCTGGGGAGCTCGGGGTGCTGTGCTGTGAGGGGCTCAGGCGATCCACGCT 2659
 Db 3369 TGAATGAGCTGGGGAGCTCGGGGTGCTGTGCTGTGAGGGGCTCAGGCGATCCACGCT 3428
 QY 2660 GCAAGCTCCAGACCTCGGGTGGGC 2685
 Db 3429 GCAAGCTCCAGACCTCGGGTGGGC 3454

RESULT 8

ACD03624

ID ACD03624 standard; cDNA; 1683 BP.

XX AC ACD03624;

XX 01-AUG-2003 (first entry)

XX Novel human GPCR related protein NOV2b cDNA.

XX Human; G-protein coupled receptor related protein; GPCR related protein;
 KW NOV; cytosolic; cardiant; antiarteriosclerotic; antidiabetic;
 KW immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic;
 KW antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine;
 KW NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer;
 KW diabetes; immune disorder; Parkinson's disease; asthma;
 KW haematopoietic disorder; Parkinson's disease; Alzheimer's disease;
 KW infection; multiple sclerosis; cancer-associated cachexia;
 KW wasting disorder; chronic disease; neurogenesis; cell differentiation;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis;
 KW chromosome mapping; tissue typing; preventive medicine; pharmacogenomic;
 KW gene; ss.

XX Homo sapiens.

XX OS W020299116-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-US017428.

XX 04-JUN-2001; 2001US-0295607P.

XX 04-JUN-2001; 2001US-0295661P.

XX 06-JUN-2001; 2001US-0296404P.

XX 06-JUN-2001; 2001US-0296418P.

XX 14-JUN-2001; 2001US-0298285P.

XX 15-JUN-2001; 2001US-0298556P.

XX 21-JUN-2001; 2001US-0299949P.

XX 28-JUN-2001; 2001US-0300883P.

XX 28-JUN-2001; 2001US-0301550P.

XX 13-AUG-2001; 2001US-0311972P.

XX 27-AUG-2001; 2001US-0315071P.

XX 29-AUG-2001; 2001US-0315660P.

XX 14-SEP-2001; 2001US-0322293P.

XX 17-SEP-2001; 2001US-0322706P.

XX 14-DEC-2001; 2001US-0341186P.

XX 28-FEB-2002; 2002US-0361189P.

XX 12-MAR-2002; 2002US-0363673P.

XX 12-MAR-2002; 2002US-0363676P.

XX 03-JUN-2002; 2002US-0363676P.

(CURA-) CURAGEN CORP.

PI Anderson DW, Baumgartner JC, Boldog FL, Casman SJ, Edinger SR;
 PI Gangolli EA, Gerlach VL, Gorman L, Guo X, Hjalt T, Kekuda R, Li L;
 PI Macdougall JR, Malyankar UM, Millet I, Padigar M, Patturajan M;
 PI Pena CE, Rastelli L, Shimkets RA, Stone DU, Spytek KA, Vernet CM;
 PI Voss EZ, Zerhusen BD;
 XX WPI; 2003-140627/13.
 DR P-PSDB; ABU9120.

XX

PT New NOVX polypeptides and nucleic acids, useful for preventing or
 PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

PS Claim 20; Page 99-100; 332pp; English.

XX The invention describes an isolated polypeptide (I) comprising any of 27
 CC 118-961 residue amino acid sequences, given in the specification, a
 CC mature form of them, a sequence that is at least 95 % identical to them,
 CC or a sequence having one or more conservative substitutions in them. The
 CC polypeptide is useful in manufacturing a medicament for treating a
 CC syndrome associated with a human disease selected from a pathology
 CC associated with the polypeptide. The NOVX polypeptides, polynucleotides
 CC and antibodies are useful in treating or preventing NOVX-associated
 CC disorders, e.g. cardiomyopathy, atherosclerosis, cancer, diabetes, immune
 CC disorders, AIDS, obesity, asthma, haematopoietic disorders, Parkinson's
 CC disease, Alzheimer's disease, infections, multiple sclerosis, cancer-
 CC associated cachexia, and other wasting disorders associated with chronic
 CC diseases. The nucleic acids and polypeptides may also be used as targets
 CC for the identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods. The nucleic acids are further used as
 CC hybridisation probes, in chromosome mapping, tissue typing, preventive
 CC medicine, and pharmacogenomics. The polypeptides are also useful as
 CC vaccines. This sequence encodes a novel human G-protein coupled receptor
 CC related protein NOV

SQ Sequence 1683 BP; 343 A; 483 C; 508 G; 349 T; 0 U; 0 Other;

Query Match 44.9%; Score 1394.8; DB 7; Length 1683;

Best Local Similarity 90.1%; Pred. No. 0;

Matches 1577; Conservative 0; Mismatches 2; Indels 171; Gaps 1;

QY 1101 CGAGCCTCTCTTACCATGTGCTTCCCTCGTGGTGTGGTGTGTACTGCGCT 1160

Db 2 CGCGCCTCTCTTACCATGTGCTTCCCTCGTGGTGTGGTGTGTACTGCGCT 61
 |||
 |||

QY 1161 CCAGCAGCAGTGGAGGGTGGGGGCTGTGAGACAGACGTCACAGGACCACTGCGAGT 1220

Db 62 CCAGCAGCAGTGGAGGGTGGGGGCTGTGAGACAGACGTCACAGGACCACTGCGAGT 121
 |||
 |||

QY 1221 GTACATGCTCTTACCTGCTGAGTCTGTGACCAACCGGGGGCCCGCCCTCCAGCC 1280

Db 122 GTACATGCTCTTACCTGCTGAGTCTGTGACCAACCGGGGGCCCGCCCTCCAGCC 181
 |||
 |||

QY 1281 CCACCCCAACAGAGAGGGTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAT 1340

Db 182 CCACCCCAACAGAGAGGGTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAT 241
 |||
 |||

QY 1341 CCTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGAGACGCTCTCGCTT 1400

Db 242 CCTATTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGAGACGCTCTCGCTT 301
 |||
 |||

QY 1401 CCTCAACATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCA 1460

Db 302 CCTCAACATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCA 361
 |||
 |||

QY 1461 CTTGAGTTTCCAGGAATCTTTTGGAGCTATGTACTATATCTTGGACGAGGGGAGGGGG 1520

Db 362 CTTGAGTTTCCAGGAATCTTTTGGAGCTATGTACTATATCTTGGACGAGGGGAGGGGG 421
 |||
 |||

QY 1521 GGAGGGCCAGACAGGAGCTGACCGGCTGTGACCGAGTACGCTTTCTGAAAGGAG 1580

Db 422 GGAGGGCCAGACAGGAGCTGACCGGCTGTGACCGAGTACGCTTTCTGAAAGGAG 481
 |||
 |||

QY 1581 CTTCTGGCACTCACAGCGCTTCTTGTGGAATCTCTTCTGAAACGAGGAGCCAGGCCA 1640

Db 482 CTTCTGGCACTCACAGCGCTTCTTGTGGAATCTCTTCTGAAACGAGGAGCCAGGCCA 541
 |||
 |||

QY 1641 CCTGGAGAGAGTCTCTGCTGGAGAGGTCTCGCCGACATCAAGATGACCTGTGTCAGTG 1700

Db 542 CCTGGAGAGAGTCTCTGCTGGAAGTCTCGCGCACATCAAGATGACCTGTGTCAGTG 601
QY 1701 GATCCAAAGCAAGCTCAGAGCGACGCTCCACCTTGCAGCGAGGCTCCTTGGAGTTC 1760
Db 602 GATCCAAAGCAAGCTCAGAGCGACGCTCCACCTTGCAGCGAGGCTCCTTGGAGTTC 661
QY 1761 CAGCTCTTTGACGAGATCCAGGAGGAGGAGTTTATCCAGCAGCGCTCAGCCACTTCCA 1820
Db 662 CAGCTCTTTGACGAGATCCAGGAGGAGGAGTTTATCCAGCAGCGCTCAGCCACTTCCA 721
QY 1821 GGTGATCGTGTGTCAGCAATTCCTTCAAGATGGAGACATGGTCTCTCTGTTCTGTCT 1880
Db 722 GGTGATCGTGTGTCAGCAATTCCTTCAAGATGGAGACATGGTCTCTCTGTTCTGTCT 781
QY 1881 GAAGCGCTCAGAGCGCGCAGGTGTGCACTTGTATGGCCACCTCAGCGCGACCG 1940
Db 782 GAAGCGCTCAGAGCGCGCAGGTGTGCACTTGTATGGCCACCTCAGCGCGACCG 841
QY 1941 GGAAGACCGCGAGGTGTCTCCGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACCAGA 2000
Db 842 GGAAGACCGCGAGGTGTCTCCGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACCAGA 901
QY 2001 GAGGACCGTCTGCTGGAGCGCTACAGTGAACATCTGGCAGCGCGCTGTGCACCAATCC 2060
Db 902 GAGGACCGTCTGCTGGAGCGCTACAGTGAACATCTGGCAGCGCGCTGTGCACCAATCC 961
QY 2061 AAACCTGATAGAGTGTCTCTGTACCGAAATGCCCTGGCAGCGCGGGGTGAAGTGTCT 2120
Db 962 AAACCTGATAGAGTGTCTCTGTACCGAAATGCCCTGGCAGCGCGGGGTGAAGTGTCT 1021
QY 2121 CTGTCAAGGACTCAGACACCCCAACTGCAAACTTCAGAACTGAGGCTGAAAGGTGCGG 2180
Db 1022 CTGTCAAGGACTCAGACACCCCAACTGCAAACTTCAGAACTGAGGCTGAAAGGTGCGG 1081
QY 2181 CATCTCCAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCCAAATGAATTTGAC 2240
Db 1082 CATCTCCAGCTCAGCTCGAGGACCTCTCTGAGCTCTCATAGCCAAATGAATTTGAC 1141
QY 2241 AAGGATGATCTCAGTGGCAACGGGGTGGATTCCAGGACATGCTCTTTGGAGGG 2300
Db 1142 AAGGATGATCTCAGTGGCAACGGGGTGGATTCCAGGACATGCTCTTTGGAGGG 1201
QY 2301 CTTGGGATCCCAAGTGCAGCTGCAGATGATTCAGTTTGAAGAAAGTGTGAGTGCAGTC 2360
Db 1202 CTTGGGATCCCAAGTGCAGCTGCAGATGATTCAGTTTGAAGAAAGTGTGAGTGCAGTC 1237
QY 2361 CGGGGCTTGTAGGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGGTTGAGTTGA 2420
Db 1238 ----- 1237
QY 2421 CTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTATGCCAGGACTGAGGCA 2480
Db 1238 ----- 1237
QY 2481 CCCAGTCTGCAGACTACGAGCTTTGTGGTGAAGATCTGCCCTCACTGCTGCTGCTG 2540
Db 1238 -----GCTGAAGATCTGCCCTCACTGCTGCTGCTGCTG 1270
QY 2541 TGACAGCTGGGCTCAACTCTCAGTGTGNAACAGAGCTTGAGAGCTGAGCCTGAGCCT 2600
Db 1271 TGACAGCTGGGCTCAACTCTCAGTGTGNAACAGAGCTTGAGAGCTGAGCCTGAGCCT 1330
QY 2601 GAATGAGCTGGGGACCTCGGGGTGCTGCTGTGTGAGGAGGCTCAGGATCCCACTG 2660
Db 1331 GAATGAGCTGGGGACCTCGGGGTGCTGCTGTGTGAGGAGGCTCAGGATCCCACTG 1390
QY 2661 CAAGCTCCAGACCTCGGGTGGGATCTGCCGCTGGGCTCTGCGGCTGTGAGGCTCT 2720
Db 1391 CAAGCTCCAGACCTCGGGTGGGATCTGCCGCTGGGCTCTGCGGCTGTGAGGCTCT 1450
QY 2721 TTCTGTGTGCTCAGGGCAACCAACCTCGGAGCTGAGCTTGAAGTTTCAACGACCT 2780
Db -----

Db 1451 TTCTGTGTGCTCCAGGCGCAACCAACACCTCCGGAGCTGGACTTGAGTTTCAACGACCT 1510
QY 2781 GGGAGACTGGGCTTGTGTTGCTGCTGAGGGCTGCAACATCCCGCTTGCAGACTCCA 2840
Db 1511 GGGAGACTGGGCTTGTGTTGCTGCTGAGGGCTGCAACATCCCGCTTGCAGACTCCA 1570
QY 2841 GAAACTGTGG 2850
Db 1571 GAAACTGTGG 1580
RESULT 9
AAD14323
ID AAD14323 standard; cDNA; 3857 BP.
AC AAD14323;
XX
DT 01-NOV-2001 (first entry)
XX
DE Human PYRIN 1 (PYRIN-1) cDNA.
XX
KW Human; nucleotide binding site; NBS; PYRIN-1; apoptotic; cytostatic;
dermatological; vasotropic; melanoma; gene therapy; ophthalmological;
cancer; leukaemia; carcinoma; arthritis; viral infection; allergy;
autoimmune disease; systemic lupus erythematosus; SLE; neutropenic; asthma;
glomerulonephritis; neurological disorder; Alzheimer's disease; ALS;
amyotrophic lateral sclerosis; Parkinson's disease; Huntington's disease;
aplastic anaemia; ischaemia; meningitis; liver disease; Crohn's disease;
insulin-dependent diabetes; multiple sclerosis; Grave's disease; HIV;
human immunodeficiency virus; tuberculosis; lepromatous leprosy; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
CDS 139..3243
FT /*tag= a
FT /note= "Human PYRIN-1 protein"
XX
PN WO200161005-A2.
XX
PD 23-AUG-2001.
XX
PF 16-FEB-2001; 2001WO-US040145.
XX
PR 17-FEB-2000; 2000US-00506067.
PR 01-SEP-2000; 2000US-00653901.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Bertin J;
XX
DR WPI; 2001-514773/56.
DR P-PSDB; AAE07514.
XX
PT Novel polypeptide comprising NBS-1 or PYRIN-1 of the pyrin domain protein
family useful in screening and detection assays and for treating, e.g.,
cancer, viral infections, autoimmune disease, and Alzheimer's.
XX
PS Claim 1; Fig 4; 11pp; English.
XX
CC The invention relates to human NBS-1 nucleotide binding site) and PYRIN-1
protein and nucleic acids. NBS-1 and PYRIN-1 DNA, protein and their
modulators are useful in the treatment of apoptotic and inflammatory
disorders, cancer (leukaemia, melanoma, carcinoma); viral infections
(including herpesvirus and adenovirus), autoimmune diseases (systemic
lupus erythematosus (SLE), immune-mediated glomerulonephritis, arthritis)
; neurological disorders (Alzheimer's disease, Parkinson's disease,
amyotrophic lateral sclerosis (ALS), retinitis pigmentosa, Huntington's
disease); aplastic anaemia, ischaemia, meningitis, liver diseases. NBS-1
and PYRIN-1 DNA, protein and their modulators are also used for the
treatment of inflammatory and immune disorders such as chronic
inflammatory diseases such as Crohn's disease, insulin-dependent
diabetes, organ specific autoimmunity, including multiple sclerosis,

CC	Grave's disease, allergy, asthma, HIV, tuberculosis and lepromatous
CC	leprosy. The present sequence is a cDNA encoding human PYRIN-1 protein
XX	
SEQ	Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;
Query Match	
Best Local Similarity 58.6%; Score 800.4; DB 4; Length 3857;	
Matches 1544; Conservative 0; Mismatches 1016; Indels 74; Gaps 6;	
QY	281 TGGTGAGGGATCCCAGAGAACCTACAGGACTATGTCGACAGAAATTCGGCTCATGG 340
DB	
QY	527 TGAAGAAGATTACCGTAAGNAGTACAGAAGTACGTGAGAAGCAGANTTCATGTCATTG 586
DB	
QY	341 AAGACCGCAATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCC 400
DB	
QY	587 AAGACAGGAATGCCGTCTGGTGAGAGTGTGAGCTCTCAACAAACGCTACACACGACTGC 646
DB	
QY	401 TGTGTGTGAGGAGCACTCAAAACCCATGAGGTTCAGACGAGCTTCTTGACACAGGCC 460
DB	
QY	647 GTCTCATCAAGGAGCAACCGAGCCAGCAGGAGAGGAGGAGCTTCTGGCCATCGGC- 705
DB	
QY	461 GGGGACACGGGAGGACCGTGGGACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTG 520
DB	
QY	706 -----AAGACCAAGACGTGTGAGAGCCCGTGAGTCCCATTAAGATGGAGTTCTGTTTG 760
DB	
QY	521 AGCCAGACGAGAGGCGGCCCGAGCCACCGCGCACCGTGTGATGCAAGAGCGCGCAGGGA 580
DB	
QY	761 ACCCCGATGATGAGCATTTCTGAGCCTGTGCACACCGTGGTGTTCAGAGGGCGCGCAGGA 820
DB	
QY	581 TAGCAAGTCCATCTGSCACACAAGGTGATCTGGACTGCGCGGAGCGGCAAGCTCTTCC 640
DB	
QY	821 TTGGAAAAACATCTCGGCCAGGAAGATGATTTGGACTGGGCATCGGGAGCACTCTACC 880
DB	
QY	641 AAGCAGATTTGATTATCTCTTCTACATCAACTCAGCGAGATGAACACGAGTGCACCG 700
DB	
QY	881 AAGACAGTTTGACTATCTGTTCTATATCCACTGTGCGGAGGTGAGCC---TTGTGACAC 937
DB	
QY	701 AATGACGATGCAAGACCTCATCTTACGTCTGCTGGCTGAGCCAGCGCGCTCTCTCAGG 760
DB	
QY	938 AGAGGAGCTTGGGGACCTGATCATGAGCTGCTGCCCGGACCCCAAAACCCCATCCACA 997
DB	
QY	761 AGCTCATCCGAGTTCCGAGCGCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGC 820
DB	
QY	998 AGATCGTGAGAAAAACCTTCAGAAATCTCTTCTCATGACGGCTTCGATGAGCTGCAAG 1057
DB	
QY	821 CTTCTTTCCAGATCTCTCAGGACCCCTGGTGCCTCTGCTGGAGGAGAAAACGGCCACCG 880
DB	
QY	1058 GTGCTTTGACGACACATAGACCGCTCTGCACCTGACTGTCAGAGCGCGAGCGGGAG 1117
DB	
QY	881 AGCTGCTCTTAAACAGCTTAATTGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCA 940
DB	
QY	1118 ACATTCCTCTGAGCAGCCTCATCAAGAAGAGCTGCTTCCCGAGGCCTCTCTGCTCATCA 1177
DB	
QY	941 CCACAGGCCCCACGCTTTTGAGAAGCTCCACCGCTCTGCTGGAGCACCCACAGGATGTGG 1000
DB	
QY	1178 CCACGAGACCTGTGGCCCTGAGAAAACCTGACAGCTTGTGGACCATCTCCGCAATGG 1237
DB	
QY	1001 AGATCTCGGGTTCTCTGAGCGCAAGAGAAAGGAATATCTTCTCAAGATATTTCCACAATG 1060
DB	
QY	1238 AGATCTCGGGTTCTCCGAGGCCAAAAGAAAAGAGTAGTACTTCTTCAAGTACTCTCTCATG 1297
DB	
QY	1061 CAGAGCAGCGGGCCCAAGTCTTCAATTACGTGAGGGACAGAGCCCTCTCTTACCAATGT 1120
DB	
QY	1298 AGGCCCAAGCCAGGCGACCTTTCAGTCTGATTCAGAGGAACAGAGGTCTCTTACCAATGT 1357
DB	
QY	1121 GCTTCGTCCTCGTGTGCTGGGTGTGTGTACTCTGCTCCACGACGACTGGAGGGGT 1180
DB	
QY	1358 GCTTCATCCCTCGTCTGTGGATCTGTTGCACTGAGTCAAAACAGCATGGAGATG 1417
DB	
QY	1181 GGGGGCTGTGAGACAGACGTCACAGGACACACATGCAAGTGTAATAGTCTTACTGCTGA 1240
DB	
QY	1418 GCAAGAGCTTGTGCCAGACATCTAAGACCAACACCGGGTGTACGTCTTCTCTCTTCCA 1477
DB	

Db 2554 AACGCCCTCGGTGACCTTCGGAATCAGACTCTCTGTGTGGACTGAAGCACCTGTTGTGTC 2613
 Qy 2320 AGGCTGCAGATGATTCAGTTTGTAGGAGTGTGACGTGAGTCCGGGCTGTGACGAGATG 2379
 Db 2614 ATCTGAAGAAGCTCTGGTTGTGTCAGTCTGCTTCCATCAGCATGTTGTTCAGGATCTT 2673
 Qy 2380 GCTTCTGTGCTCGGCACCAACCCACACTCTGTTGTGAGTTGGACCTGACAGGAATGCACTG 2439
 Db 2674 GCATCAGTATTGAGCACCAGCCATTCCTTGACAGACTCTATGTGGGGAGATGCTTGT 2733
 Qy 2440 GAGGATTTGGGCTGAGTTTACTATCCAGGGAGTCCAGGACCCAGTCTGCGACTACGG 2499
 Db 2734 GGAGACTCAGGAGTGCAGATTTTATGTGAAAAAGCAAGATCCACAGTGTAACTTCGAG 2793
 Qy 2500 ACTTTGTGCTGAAGATCTCGGCTCCTCACTGCTGCTGCTGCTGACAGCTGCTCACT 2559
 Db 2794 AAACCTGGGTGTGTAATCTTGGCCCTTACGTACGTCTGTGTGCTGTGCTTGTCTCGGTA 2853
 Qy 2560 CTCAGTGTGAACAGAGCTGAGAGCTGGAGCTGAGCTGAGCTGAAATGAGCTGGGGACCTC 2619
 Db 2854 CTCAGCACTAATCAGAACTCTCAGCACCCTTTACCTCGAGGCAACACTCTCGAGACAAG 2913
 Qy 2620 GGGGTCTGCTGTGTGAGGGCTCAGGATCCACAGTCCAGGCTCCAGACCTTCGGG 2679
 Db 2914 GGGATCAAACTACTCTGTGAGGAGCTCTTGCACCCGAGCTGCAAGCTTCAGGTGTTGGAA 2973
 Qy 2680 TTGGGCTGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2739
 Db 2974 TTAGCACTGCACTCAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3033
 Qy 2740 AACCAACCTCCGGAGCTGAGCTTGTGCTTCAAGCACTGGGAGCTGGGGCTGTGG 2799
 Db 3034 AGCCAGAGCTCGGAAGCTGAGCTGGGCAACATGACTGGGCACTGGGGCTGATG 3093
 Qy 2800 TTGCTGGCTGAGGGCTGCAATCCCGCTGCGAGACTCCAGAACTGTGGCTG 2853
 Db 3094 ATGTTCTGTGAGTGTGTAACAGCAGAGCTGCTCTCTGCAAGAACCTGGGGTTG 3147

RESULT 10

ABX93556

ID ABX93556 standard; cDNA; 3857 BP.

XX AC ABX93556;

XX AC ABX93556;

DT 22-MAY-2003 (first entry)

XX Huma cDNA encoding PYRIN-1.

XX Huma; ss; gene; nucleotide binding site; pyrin domain; NBS-1; PYRIN-1;

XX caspase-1; anti-inflammatory; apoptosis; ASC; NF-kB; nuclear factor KB;
 XX LRR; leucine rich repeat; inflammatory disorder; familial cold urticaria;
 XX arthritis; inflammatory bowel disease; Crohn's disease; asthma; allergy;
 XX ulcerative colitis; rheumatoid arthritis; Lyme disease; Grave's disease;
 XX insulin-dependent diabetes; multiple sclerosis; contact dermatitis;
 XX psoriasis; graft versus host disease; food allergy; conjunctivitis;
 XX chronic obstructive pulmonary disease; food allergy; conjunctivitis;
 XX helminthic infection; Leishmaniasis; viral infection; HIV infection;
 XX bacterial infection; tuberculosis; leprosy; chromosome 1q44.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 139..3243

FT /tag= a

FT /product= "PYRIN-1"

XX US2002187922-A1.

XX PD 12-DEC-2002.

XX PD 12-DEC-2002; 2002US-00127516.

XX 17-FEB-2000; 2000US-00506067.
 PR 01-SEP-2000; 2000US-00653901.
 PR 26-SEP-2001; 2001US-00964955.
 PR 20-DEC-2001; 2001US-00027629.
 XX (BERT/) BERTIN J.
 PA (MANJ/) MANJI G A.
 PI Bertin J, Manji GA;
 XX WPI; 2003-328763/31.
 DR P-PSDB; ABU08503.
 XX Identification of compound that binds to polypeptide (for e.g. a PYRIN
 PT protein), useful for treating disorders associated with inappropriate
 PT apoptosis, for e.g. inflammatory disorder.
 XX Disclosure; Fig 4; 74pp; English.

XX The invention relates to a new method for the identification of a
 CC compound that binds to a human PYRIN-1 (an apoptotic signalling molecule
 CC appearing as ABU08503) comprising contacting the polypeptide or a cell
 CC expressing the polypeptide to a test compound and determining whether
 CC PYRIN-1 binds to the test compound, or determining the effect of the test
 CC compound on the activity of PYRIN-1. In a similar manner, compounds are
 CC isolated which modulate the binding of PYRIN-1 to ASC (comprising testing
 CC the compounds against the pyrin binding domains of both PYRIN-1 and ASC
 CC (not defined)), modulate the activity of NF-kB (nuclear factor KB), and
 CC which modulate the ASC-mediated activation of NF-kB (comprising: (a)
 CC measuring the binding of a test compound to the LRR (leucine rich repeat)
 CC domain of PYRIN-1; and (b) measuring the activation of NF-kB in a cell
 CC expressing ASC and PYRIN-1 in the presence and absence of the LRR domain
 CC binding compound). The candidate modulators may be useful for treating an
 CC inflammatory disorder. Also included is a method (M8) for identifying a
 CC modulator of caspase-1 activity. The identified compound can be used to
 CC modulate the polypeptide's activity (ASC and NF-kB activities in a
 CC patient). By modulating the expression or activity of the polypeptide
 CC (PYRIN-1 and ASC), a disorder associated with inappropriate apoptosis
 CC (inflammatory disorders e.g. familial cold urticaria, arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC rheumatoid arthritis, Lyme disease, insulin-dependent diabetes, multiple
 CC sclerosis, Grave's disease, contact dermatitis, psoriasis, graft versus
 CC host disease, asthma, chronic obstructive pulmonary disease, allergies
 CC (e.g. food allergies), conjunctivitis, helminthic infection (e.g.
 CC Leishmaniasis), viral infections such as HIV infection, and bacterial
 CC infections such as tuberculosis and leprosy) can be treated. Human NBS-1
 CC protein (nucleotide binding site protein 1) contains a pyrin domain which
 CC was used to identify genomic sequences containing the PYRIN-1 gene. The
 CC gene for pyrin-1 is located on chromosome 1q44. The present sequence
 CC encodes human PYRIN-1

SQ Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;

Query Match 25.8%; Score 800.4; DB 8; Length 3857;
 Best Local Similarity 58.6%; Pred. No. 4.3e-188;
 Matches 1544; Conservative 0; Mismatches 1016; Indels 74; Gaps 6;

Qy 281 TGTGTGAGGATCCCGAGGAACCTACAGGAGCTATGTCCGAGGAAATTCGGCTCATGG 340
 Db 527 TGAAGAAAGATTACCGTTAAGAGTACAGAAAGTACGTGAGAGCAGATTCCAGTGCATTG 586
 Qy 341 AAGACCGCAATGCGGCGCTAGGGGAATGTGTCAACCTCAGGCACCGGTACACCGGCTCC 400
 Db 587 AAGACAGGAATGCGGCTCTGGGTGAGAGTGTGAGCTTCAACAAAGCTACACAGACTGC 646
 Qy 401 TGCTGGTGAAGGAGCAGCTCAAAACCCATCGAGTCCAGCAGAGCTTCTGGACACAGGCC 460
 Db 647 GTCTCATCAAGAGAGCAGCGGAGCCAGCAGGAGGAGGAGAGGAGCTTCTGGCATCGGC- 705
 Qy 461 GGGGACACGCGGAGGAGCAGCTGGGACACACAGGCTAGGCCCATCAAGATAGAGACCTCTTTG 520
 Db 706 -----AAGACCAAGACGTTGTGAGAGCCCGCTGAGTCCCATTAAGATGAGTGTCTTTG 760

Qy 521 AGCCAGAGGAGCGCCCGAGCCACCGGACACCGTGGTCAATCAAGGCGCGCAGGGA 580
Db |||||
Qy 761 ACCCGATGATGACATCTGAGCTGTGACACCGTGGTGTTCACGGGCGCGCAGGGA 820
Db |||||
Qy 581 TAGCAAGTCCATGCTGGCACAAAGGTGATGCTGGACTGGCGGACGGGAAGCTCTTCC 640
Db |||||
Qy 821 TTGGGAAACAACTCTGCGCCAGGAGATGATGTTGGACTGGGCATCGGGGACACTCTACC 880
Qy 641 AAGCAGATTGATATCTCTTCTACATCAACTGCAGGAGATGAACACAGAGTGCACGG 700
Db |||||
Qy 881 AAGACAGGTTTGACTATCTGTTCTATATCCACTGTGGGAGTGAGCC---TTGTGACAC 937
Qy 701 AATCAGAGTCAAGAACCTCACTTCTCAAGTCTGCTGGCTGAGCCACCGGCGCTCTCCAGG 760
Db |||||
Qy 938 AAGAGAGCCTGGGGACCTGATCATGATGACTGCTGCCGACCCAAACCCACCCATCCACA 997
Qy 761 AGCTCATCCGAGTTCGAGCGCCCTCTTTTCATCATGACGGCTTGATGAGCTCAAGC 820
Db |||||
Qy 998 AGATCGTGAGAAAACCTCCAGAACTCTCTTCTCATGAGACGGCTTGATGAGCTGCAAG 1057
Qy 821 CTTCTTTCCAGATCTCTCAGGAGCCCTGGTGGCTCTGCTGGGAGGAGAAACGGGCCACGG 880
Db |||||
Qy 1058 GTGCTTTGAGGACACATAGACCGCTCTGCACTGACTGACAGAGCCCGGCGGAG 1117
Qy 881 AGCTGCTTTTAAAGCTTAATTCGGAAGAAGTGTCTCCCTGAGCTATCTTTGCTCATCA 940
Db |||||
Qy 1118 ACAITCTCCTGAGCAGCCTCATCAAGAAAGAGTGTCTCCGAGGCGCTCTGCTCATCA 1177
Qy 941 CCACAGCGCCACCGCTTTGAGAGAGTCCACCGTCTGCTGGAGCACCACCGGATGCG 1000
Db |||||
Qy 1178 CCACAGAGCCTGTGGCCCTGAGAAACTGCAAGACTTGTGGACCATCTCCGCGATGCG 1237
Qy 1001 AGATCTGGGCTTCTCTGAGGAGAAAGGAGAACTCTCAAGATATTTCCACAAATG 1060
Db |||||
Qy 1238 AGATCTGGGTTTCTCCAGGCGCAAGAGAGTACTTCTTCAAGTACTTCTCTGATG 1297
Qy 1061 CAGAGCAGCGGGGCGAAGTCTTCAATACGTGAGGAGCAAGAGCCTCTCTTCAACATGT 1120
Db |||||
Qy 1298 AGGCCAAAGCAGGCGAGCCTTCACTGATTCAGGAGAACGAGTCTCTTCAACATGT 1357
Qy 1121 GCTTGTCTCCCTGGTGTGCTGGGTGTGTGTAACCTGCTCCAGCAGCAGCTGGAGGTG 1180
Db |||||
Qy 1358 GCTTCACTCCCTGCTGCTGGATCGGTGCACTGGAGTGAACAGCAGAGTGGAGAGTG 1417
Qy 1181 GGGGCTGTTGAGACAGAGCTCCAGGACCACTGCAAGTGTACATGCTCTACCTGTGA 1240
Db |||||
Qy 1418 GCAAGAGCCTTGCCAGACATCTAAGACCAACACCGCGGTGTACGTCTTCTTCTTCCA 1477
Qy 1241 GTCTGATGCAACCAAGCCGGGGGCGCCGCGCTCCAGCCGCCACCCAAACAGAGAGGT 1300
Db |||||
Qy 1478 GTTTGCTGCAAGCCCGGGAGGAGCCAGGAGCAAGGCTCTGCGCCACCTCTGGGGGC 1537
Qy 1301 TGTGCTCTTGGCGGAGATCGGCTCTGGAATCAGAAATCTATTTTTCAGGAGCAGGACC 1360
Db |||||
Qy 1538 TCTGCTCTTGGCTGCAATGGAATCTGGAACCAAGAAATCTGTTGAGGAGTCCGACC 1597
Qy 1361 TCCGGAAGCAGCGCTAGACCGGGAAGACGCTCTGCGCTTCTCAACATGAACATCTTCC 1420
Db |||||
Qy 1598 TCAGGAATCATGGAATGCAAGAGCGGATGTCTGCTTTCTGAGGATGAACCTGTTC 1657
Qy 1421 AAGAGGACATCAACTGTGAGAGGTACTACAGCTTCACTCACTGAGTTTTCAGGAATCT 1480
Db |||||
Qy 1658 AAAAGGAAGTGGACTGCGAGAAGTTCTACAGCTTCACTCAATGACTTTTCCAGGAGTCT 1717
Qy 1481 TTGCAAGTATGTAATCTTGAAGAGGAGGAGGCGG----- 1520
Db |||||
Qy 1718 TTGCGCATGTAATCTTCTGGAAGAGGAGAAAGAAAGAGAGGAGGAGGAGGAGGAGG 1777
Qy 1521 -----GGCAGGCGCCAGACAGGAGCGTGACAGGCTGTGACCGAGTACGCGTTTCTG 1573
Db |||||
Qy 1778 GTCGTTTGAAGCTTCCAGCGGAGAGCGTGACAGTCTTCTGGAAGAACTATGCAAAATTCG 1837

Qy 1574 AAAGAGGCTTCTTGGACTCAGCAGCGCTTCCCTGTTTGGACTCCTGAAACAGGAGACCA 1633
Db |||||
Qy 1838 AAAAGGGGTATTTGATTTTGTGTTTGTGTTTCTCTTTGGCTGTTTAAACAGGAGGGA 1897
Db |||||
Qy 1634 GGAGCCACCTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGACATCAAGATGACCTGT 1693
Db |||||
Qy 1898 CCTCCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGAAATCAGGCTGAGCTGC 1957
Qy 1694 TGCAAGTGCATCAAGCAAGCTCAGAGCGAGCTCCACCCTCAGCAGGCGCTCTTGG 1753
Db |||||
Qy 1958 TGAATGATGTAAGTGAAGCAAGCAAGCTAAAGCTGACAGATCCAGCCAGCAGCTGG 2017
Qy 1754 AGTCTTTCAGCTGCTGTAACAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTGAGCC 1813
Db |||||
Qy 2018 AATTGTTCTACTGTTGTACAGATGCAAGAGGAGGACTTCTGTCGAAAGGCGCATGGA 2077
Qy 1814 ACTTCCAGGCTGATCGTGTGAGCAACATGCTCCAGATGAGCAGATGCTCTCTCGT 1873
Db |||||
Qy 2078 ATTTCCCAAGTTGAG---ATCAATCTCTCCACAGATGAGCAGATGATGTTTCTCT 2134
Qy 1874 TCTGTCTGAAGCGCTGAGGAGCGCCAGGTGCTGCACTTGTATGCGCGCCACCTACAGG 1933
Db |||||
Qy 2135 TTTGATTTGAGAACTGTCTCATCGGTGGAGTCACTGCTCCTGGGTTTCTCCATAACATGC 2194
Qy 1934 CGAGCGGGGAACA-----CCGCGCGAGGTGC 1959
Db |||||
Qy 2195 CCAAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2254
Qy 1960 TCCGAGGAGCGCACAGCTGCTGCTGAGCTCAGACCAGAGAGACCGTCTCTGCTGGAC 2019
Db |||||
Qy 2255 TCCCAAGCTCTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2313
Qy 2020 GCCTACAGTGAACATCTGGCAGCGCCCTGTGCAACCAATCCAAACCTGATAGAGCTGT 2079
Db |||||
Qy 2314 AGTTTTTCCGCGGCGCTCTTTTTCAGTTCTGAGCACCAGCAGAGTCTAACTGAATGAGC 2373
Qy 2080 CTGTACCGAAATGCTGGGCGAGCGGGGCTGAGCTGCTCTGTCAGGAGTCAAGAC 2139
Db |||||
Qy 2374 CTCAAGTGAACATCTCTGGGGGAGCCAGGAGTGAAGTGTGTGTGAACCTCCAGAT 2433
Qy 2140 CCCAACTCAAACTTTCAGAACTGAGGCTGAGAGGTGCGCATCTCCAGCTCAGCCTGC 2199
Db |||||
Qy 2434 CTGCTGTGAACATCTGGAGATTTGTGTTGGGCGCTGTGGCTCTCGCATGAGTGTGCT 2493
Qy 2200 GAGGACCTCTCTGAGCTCTCATAGCCAAATGAAGATTTGCAAGAGTGAATCTCAGTGGC 2259
Db |||||
Qy 2494 TTCGACATCTCTTGGTCTCTCAGCAGCAACCCAGAGCTGGTGGAGCTGGACCTGAGTGC 2553
Qy 2260 AACGCGTTGGATTTCCAGGCAATGATGCTGCTTTGCGAGGCGCTGCGCATCCCGAGTGC 2319
Db |||||
Qy 2554 AACGCTTCCGTTGACTTCCGAAATCAGACTTCTGTGTGGGACTGAAGCACTGTTGTGC 2613
Qy 2320 AGGCTGAGATGATTCAGTTGAGGAGTGTGAGTGGAGTCCGGGCTTGTTCAGGAGATG 2379
Db |||||
Qy 2614 AATCTGAAGAGCTCTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2673
Qy 2380 GCTTCTGTGCTGGCACCAACCAATCTGTTGAGTTGAGCTGAGCAGGAAATGCACTG 2439
Db |||||
Qy 2674 GCATCAGTATTGAGCACCGACCAATCTCCCTGACCACTCTATGTGGGGAGAAATGCTTGT 2733
Qy 2440 GAGGATTTGGGCGCTGAGTTTACTATGCGAGGAGTGAAGGAGGAGGAGGAGGAGGAG 2499
Db |||||
Qy 2734 GGAGACTCAGGAGTGGCAATTTTATGTGAAAGGCAAGATCCACAGTGTAACTTCGAG 2793
Qy 2500 ACTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2559
Db |||||
Qy 2794 AAACTGGGTTGGTGAATTTCTGGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2853
Qy 2560 CTCAGTGTGAACCGAGGCTGAGAGCTGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2619
Db |||||
Qy 2854 CTCAGCACTAATCAGAACTCTCAGCACCTTTTACCTCGGAGGCAACACTCTCGGAGACA 2913
Qy 2620 GGGGTGCTGCTGCTGTGAGGCGCTCAGGCACTCCACGTGCAAGCTCCAGACCTCGCG 2679

2914 GGGATCAAACTACTCTGTGAGGAGCTCTTGACCCCGAGCTCAAGCTTCAGTGTGGAA 2973
 2680 TTGGCATCTGCCGCTGGGCTCTCCGCTGTGAGGCTCTTCTGTGTGCTCCAGGCC 2739
 2974 TTAGACAACCTGCAACCTCACGCTCAACCTGCTGCTGGATCTTTCCACACTTCTGACCTCC 3033
 2740 AACCACAACCTCCGGAGCTGAGCTTGTAGTTTCAACGACCTGGGAGACTGGGGCTGTGG 2799
 3034 AGCCAGAGCTTGGAAAGCTGAGCTTGGGCAACATGATCTGGGCACTCTGGGGTCAATG 3093
 2800 TTGCTGGCTGAGGGCTGCAACATCCCGCTGCAGACTCCAGAAACTGTGGCTG 2853
 3094 ATGTTCTGTGAAGTCTGAAACAGCAGAGCTGCTCTCTGCGAGAACTGGGGTTG 3147

RESULT 11

ACD27909
ID ACD27909 standard; cDNA; 3857 BP.

XX ACD27909;

XX 23-SEP-2003 (first entry)

XX Human pyrin domain family protein PYRIN-1 cDNA.

XX Human; ss: gene; pyrin domain family; PYRIN-1; cellular differentiation;
 KW tissue typing; cellular proliferation; cell survival; apoptosis disorder;
 KW inflammatory disorder; apoptosis associated speck like protein; ASC;
 KW nuclear factor kappaB; NF-kappaB.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 139..3243
 FT /*tag= a
 FT /product= "PYRIN-1"

XX US2002197660-A1.

XX 26-DEC-2002.

XX 20-DEC-2001; 2001US-00027629.

XX 17-FEB-2000; 2000US-00506067.

XX 01-SEP-2000; 2000US-00653901.

XX 26-SEP-2001; 2001US-00964955.

XX (BERT/) BERTIN J.
 XX (MANJ/) MANJI G A.

XX Bertin J, Manji GA;

XX WPI; 2003-521543/49.

XX P-PSDB; AB063315.

XX Identifying compound that binds to polypeptide useful in treating
 PT apoptotic and inflammatory disorders, comprises contacting polypeptide
 PT with test compound and determining whether polypeptide binds to test
 PT compound.

XX Disclosure; Fig 4; 71pp; English.

XX The invention relates to a compound that binds to a polypeptide (e.g.
 CC PYRIN-1) or NBS-1. The nucleic acid molecules, proteins, protein
 CC homologues and antibodies incorporated in the invention can be used in
 CC screening assays; detection assays (e.g. chromosomal mapping, tissue
 CC typing, forensic biology); predictive medicine (e.g. diagnostic assays,
 CC prognostic assays, monitoring clinical trials, and pharmacogenomics); and
 CC methods of treatments (e.g. therapeutic and prophylactic). When NBS-1 or
 CC PYRIN-1 protein interacts with other cellular proteins, they can be used
 CC for regulation of cellular proliferation and differentiation; and
 CC regulation of cell survival. The isolated nucleic acid molecules of the

CC invention can be used to express NBS-1 or PYRIN-1 protein to detect NBS-1
 CC or PYRIN-1 mRNA or a genetic lesion in NBS-1 or PYRIN-1 gene and to
 CC modulate NBS-1 or PYRIN-1 activity. These proteins can be used to screen
 CC drugs or compounds which modulate the NBS-1 or PYRIN-1 activity or
 CC expression to treat disorders characterized by insufficient or excessive
 CC production of these proteins. The anti-NBS-1 or PYRIN-1 antibodies of the
 CC invention can be used to detect and isolate NBS-1 or PYRIN-1 proteins and
 CC modulate NBS-1 or PYRIN-1 activity. With the invention, disorders
 CC associated with inappropriate apoptosis and inflammatory disorder can be
 CC treated; and the apoptosis associated speck like protein (ASC) and
 CC nuclear factor (NF)-kappaB activity in a patient can be modulated by
 CC administering the compound that alters the activity of PYRIN-1. The
 CC method takes advantage of the interactions between NBS-1 or PYRIN-1 and
 CC various proteins involved in apoptotic and inflammatory signaling
 CC pathways. The nucleic acid molecules of the invention can be inserted
 CC into vectors. The present sequence represents cDNA encoding the human
 CC pyrin domain family protein PYRIN-1
 XX

SQ Sequence 3857 BP; 958 A; 931 C; 1023 G; 945 T; 0 U; 0 Other;

Query Match 25.8%; Score 800.4; DB 8; Length 3857;
 Best Local Similarity 58.6%; Fred. No. 4.3e-188;
 Matches 1544; Conservative 0; Mismatches 1016; Indels 74; Gaps 6;

QY 281 TGGTGGGGATCCCCAGGAACCTACAGGAGTATGTCCGAGGAAATTCGGCTCATGG 340
 DB 527 TGAAGAAAGATTACCGTAAGAGTACAGAAAGTACGTGAGAGCAGATTCAGTGCATGTG 586
 QY 341 AAGACCGCAATGCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCGCGTCC 400
 DB 587 AAGACAGGAATGCCCGTCTGGGTGAGAGTGTGAGGCTCAACAAACGCTACACAGCTGC 646
 QY 401 TGCTGGTGAAGGAGCAGCTCAAAACCCATGCGAGTCCAGCAGCGCTTCTGGACACAGCC 460
 DB 647 GTCTCATCAAGAGGACACCGGAGCCAGCAGGAGAGGAGAGGAGCTTCTGGCCATCGGC- 705
 QY 461 GGGACACCGGAGGAGCCGTGGGACACACAGGCTAGCGCCCATCAAGATAGAGACCTCTTTG 520
 DB 706 -----AAGACCAGAGAGCTGTGAGAGCCCGCTGAGTCCCATTAAGATGAGTTGCTTTG 760
 QY 521 AGCCAGACGAGGAGCGCCCGAGCCACCGCGCACCGTGTATGCAAGCGCGGAGGGA 580
 DB 761 ACCCGGATGATGAGCATCTGAGCGCTGTGACACCGTGTGTCACCGGTGTCAGGGGCGGAGGA 820
 QY 581 TAGGCAAGTCCATGCTGGGACACAGAGTGTGAGTGGGCGGAGCGGAGAGCTTTCC 640
 DB 821 TTGGGAAAAACAATCTCTGGCCAGGAAGATGATGTGAGCTGGGCACTCGGGGACACTCTACC 880
 QY 641 AAGCAGATTTGATTATCTTCTATCATCAACTGCGAGGAGATGAACAGAGTGCACCG 700
 DB 881 AAGACAGGTTGACTATCTGTTCTATATCCACTGTGCGGAGGTGAGCC---TTGTGACAC 937
 QY 701 AATGAGCATGCAAGACCTCATCTTTAGCTGTGCGCTGAGCCCGAGCGCGCTCTCCAGG 760
 DB 938 AGAGGAGCTGGGGAGCCCTGATCATGAGCTGTGCCCCGACCCAAACCCACCATCCACA 997
 QY 761 AGCTCATCGAGTTCGAGGCGCTCTTTTCATCATGACGGCTTCAGTGTGAGCTCAAGC 820
 DB 998 AGATCGTGAGAAAAACCCCTCCAGAAATCTCTTCTCATGAGAGGCTTCGATGATGCAAG 1057
 QY 821 CTTCTTTCCAGATCCTCAGGAGACCTCTGCTGCTGGGAGGAGAAACCGCCACCG 880
 DB 1058 GTGCTTTGACGAGCACATAGACCGCTCTGCACTGAGTGGCAGAAAGCCGAGCGGGAG 1117
 QY 881 AGCTGCTTCTTAACAGCTTAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCA 940
 DB 1118 ACATTCTCTGAGCAGCTCATCAAGAAAGAGTCTTCTCCGAGGCTCTCTGCTCATCA 1177
 QY 941 CCACAGGCCCAACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCAGGCAATGG 1000
 DB 1178 CCAGAGACCTGTGGGCGCTGGAGAACTGACAGCACTTTGCTGGACCATCTCTCGGCATGTGG 1237
 QY 1001 AGATCTGGGCTTCTCTGAGGAGGAGAAAGGAATACTTCTTACAAGTATTTTCCACAATG 1060

cytokine receptor signaling caspase-mediated proteolysis;
 c-Jun N-terminal kinase activation; cell life; cell death; apoptosis;
 inflammation; cell adhesion; cancer; keratinocyte; hyperplasia;
 neoplasia; keloid benign prostatic hypertrophy; inflammatory hyperplasia;
 fibrosis; smooth muscle cell proliferation; balloon angioplasty;
 restenosis; leukaemia; lymphoma; inflammatory disease; allergy;
 arthritis; lupus; schrojen's syndrome; Crohn's disease;
 ulcerative colitis; graft versus host disease; stroke; heart failure;
 neurodegenerative disease; parkinson's disease; Alzheimer's disease; HIV;
 cancer therapy; PAAD domain family; human; pyrin 2; gene; ss;
 PAAD domain.

Homo sapiens.

US2003077699-A1.

24-APR-2003.

25-SEP-2001; 2001US-00965621.

26-SEP-2000; 2000US-00671760.

26-SEP-2000; 2000US-0367367P.

(REED/) REED J C.

(GODZ/) GODZIK A.

(CHUZ/) CHU Z.

(PAWL/) PAWLOWSKI K.

(FIOR/) FIORENTINO L.

(ARIZ/) ARIZA M E.

(STEH/) STEHLIK C.

Reed JC, Godzik A, Chu Z, Pawlowski K, Fiorentino L, Ariza ME;

Stehlik C;

WPI: 2002-471256/50.

P-PSDB; ADE36452.

Novel isolated PAAD domain containing polypeptide useful for inducing apoptosis by inhibiting nuclear factor kappa B activation and in gene therapy for treating cancer.

Disclosure; SEQ ID NO 58; 93pp; English.

The invention describes an isolated PAAD domain containing polypeptide (I) comprising 80% identity to the amino acid sequence of PAAD and nucleotide binding protein (PAN) 2-6, pyrin 2, apoptosis-associated speck -like protein containing a caspase recruitment domain (ASC)-2 fully defined in specification, where (I) is biologically active. (I) is useful for identifying a (I)-associated polypeptide, an agent altering that association and agents that modulate PAAD domain mediated inhibition of nuclear factor kappa B (NFkappaB). A NB-ARC domain polypeptide is useful for identifying an agent that modulates the activity of the NB-ARC domain of (I). (I) or its functional fragments is useful in altering cellular or biochemical process such as apoptosis, NFkappaB induction, cytokine processing, cytokine receptor signaling caspase-mediated proteolysis or c -Jun N-terminal kinase activation, thus having modulating effect on cell life and death (apoptosis) inflammation, cell adhesion or other cellular or biochemical processes. (I) is useful for treating cancer pathologies, keratinocyte, hyperplasia, neoplasia, keloid benign prostatic hypertrophy, inflammatory hyperplasia, fibrosis, smooth muscle cell proliferation in arteries following balloon angioplasty (restenosis), leukaemia, lymphomas; inflammatory diseases such as allergies, arthritis, lupus, schrojen's syndrome, Crohn's disease and ulcerative colitis, graft versus host disease, stroke, heart failure, neurodegenerative diseases such as parkinson's and Alzheimer's disease, human immunodeficiency virus infection (HIV). (I) is useful for diagnosing cancer or monitoring cancer therapy. This sequence encodes a human PAAD domain associated protein.

Sequence 2524 BP; 590 A; 580 C; 637 G; 564 T; 0 U; 153 Other;

Query Match

Best Local Similarity 20.6%; Score 639; DB 6; Length 2524;

Matches 1069; Conservative 0; Mismatches 585; Indels 39; Gaps 4;

QY	281	TG	GTGAGGATCC	CCAGGAACCT	CAGGACTAT	GTCCGAGGAAT	TCCGGCTCAT	GG 340
DB	570	TGA	GAAGAAGATT	ACCGTAAGA	AGTACAGAA	AGTACGTGAGA	AGCAGATTCC	AGTATTC 629
QY	341	AAG	ACCGCAATCG	CGCCCTAG	GGGAATGT	GTCAACCT	CAGCCACCG	GTACACCCGGCTCC 400
DB	630	AAG	CAGGAATGCC	CGTCTGGG	TGAGAGTGT	GAGCCTCA	CAACACCT	TACACACGACTGC 689
QY	401	TGT	TGTGAAGAG	CACACTCA	AACCCCAT	TCAGGTC	CAGCAGCTT	CTTGGACACAGGCC 460
DB	690	GT	CTCATCAAG	GAGCACCG	GAGCCAG	CAGAGAGG	GAGGAGGAG	CTTCTGGCCATCGGC 748
QY	461	GG	GGACACCG	GAGGACCG	GTGGGAC	ACCCAG	CTAGCCCC	ATCAGATAGAGACCTCTTTG 520
DB	749	----	AAG	ACCAAG	AGTGTG	AGAGCC	CGCGTGA	GTGAGTGGAGTTGCTGTTG 803
QY	521	AG	CCAGAGAG	GAGCGCC	CCGAGCC	ACCGCG	CACCGTGG	TGTCATGCAAGCGCGGCAAGGA 580
DB	804	AC	CCCGATG	ATGAGCA	TTCTG	AGCCCT	GTGCAC	ACCGTGTGTTCCAGGGGGCGGCAAGGA 863
QY	581	TAG	GCAAGTCA	TGCTGG	CACACA	AGGTG	ATGCTGG	CGCGACGGGAAAGCTTTCC 640
DB	864	TT	GGAAAA	CAATCT	CTGGCC	AGGAAG	ATGATGT	TGGACTGGCGCTCGGGGACACTCTACC 923
QY	641	AAG	GCAAGT	TTGATT	TATCT	TTTAC	ATCACT	GCAGGGAGATGACACAGAGTGCCACCG 700
DB	924	AAG	CAGGTT	GTGACT	TATCT	TTTAT	CTATCT	GTGAGAGGTGAGCC---TTGTGACAC 980
QY	701	AAT	GACGAT	CAAGAC	CTCAT	CTT	CAGCT	GTGGCTTGAGCCCGCGGCTCTCCAGG 760
DB	981	AG	AGAGAG	CCCTGG	GAGACCT	GATCAT	GAGCTG	CTGCCCGACCCCAACCCACCATCCACA 1040
QY	761	AG	CTCATCG	AGTTCC	CGAGG	CCCTCT	TTTCA	TATCACTGAGCGGCTTCGATGAGCTCAAGC 820
DB	1041	AG	ATCGT	GAGAAA	ACCCTC	CAGA	ATCCT	CTTCTCATGAGCGCTTCGATGAGCTGCAAG 1100
QY	821	CT	CTCTT	CCACCAT	CTCAG	GGACCT	CGTG	CTCTGCTGGGAGGAGAAACCGCCACCG 880
DB	1101	GT	GCCTTT	GACGAC	ACATAG	GACCG	CTCTG	CATGAGTCGAGGAGGCGGCGGAG 1160
QY	881	AG	CTGCTT	CTTAA	CAGCTT	TAAT	TCGAA	GAAGCTGCTCCCTGAGCTATCTTTGCTCATCA 940
DB	1161	AC	ATTCTCT	CTGAG	CAGCTC	ATC	AGAA	AGAGCTGTTCCGAGGCTCTCTGCTCATCA 1220
QY	941	CA	CACGGCC	CCAGG	CTTTG	GAAGAG	CTCC	ACCGTCTGCTGAGCACCCAGGCAATGG 1000
DB	1221	CC	ACGAGAC	CTGTGG	CCCTGG	AGAA	AACTG	CACTTGTGACCATCTCTCGCATGTGG 1280
QY	1001	AG	ATCTCT	GGGCTT	CTCTG	AGG	CAGAA	AGGAATCTTCTACAGTATTTCCCAATG 1060
DB	1281	AG	ATCTGG	GTCTT	CCGAG	GGCC	AAA	AGAAAGATCTTCTTCAAGTACTTCTCTGATG 1340
QY	1061	C	AGAGCG	GGGCG	CAAGT	CTT	CAAT	TACGTGAGGCAACAGAGCTCTCTTCAACATGT 1120
DB	1341	AG	GGCCAA	AGCCAG	GGCAG	CCCTT	CAGT	CTGATTCAGAGNACGAGGTCTCTTCAATGT 1400
QY	1121	G	TTTCTG	CTCCCT	TGTTG	TGCTGG	TGTG	TAACCTGCTTCTGAGTGTGAGGGTG 1180
DB	1401	G	TTTCTAT	CCCCCT	TGTTG	TGCTGG	TGTG	TAACCTGCTTCTGAGTGTGAGGGTG 1460
QY	1181	G	GGGGCT	CTT	TGAG	CAGAG	CTCC	AGGACCACTGAGTGTGATGCTTCTTCTGCTGA 1240
DB	1461	G	GAAGAG	CTTTGG	CC	CAGAC	ATCTCA	AGACCAACCCAGAGAGGGGT 1520
QY	1241	GT	CTGAT	CAAC	CCCC	GGGG	CCCC	CGGCTCTCAGCCCCACCCAGAGAGGGGT 1300
DB	1521	G	TTTGT	CTG	AC	CGG	GGG	AGGACCACTGCGGCCACCTCTGGGGGC 1580
QY	1301	T	GTGCT	CTCTG	GG	CG	CAG	ATGAGTCTCTGGAATCAGAAAAATCTATTGAGGAGGAGACC 1360
DB	1581	T	CTGCT	CTT	GG	CT	GC	AGATGGAATCTGGAACCCAGAAAAATCTCTGTTGAGGAGTCCGACC 1640

Db 488 CACCTCTGGGGCTCTGCTCTTTGGCTGCAGATGGAATCTGGAACAGAAATCTGTTT 547
Qy 1348 GAGAGCAGGACCTCCGGAGCAGCGGCTAGACGGGGAAGAGCTCTGCTGCTTCCCTCAAC 1407
Db 548 GAGGAGTCCGACCTCAGGAATCATGGAATGAGAGGCGGATGTCTGCTTCTCTGAGG 607
Qy 1408 ATGAACATCTTCCAGAGGACATCACTGTGAGAGGTACTACAGTTCATCCATCTGAGT 1467
Db 608 ATGAACCTGTCTCCAAAGGAAGTGGACTGCGAGAAGTCTTACAGCTTCATCCATGACT 667
Qy 1468 TTCCAGGAATCTTTGAGAGTATGTAATATCTGGAACGAGGGGAGGGCGG----- 1520
Db 668 TTCCAGGAGTCTTTGCGGCATGTACTCTGCTGGAAGGAAGGAAGGAGGAGGAG 727
Qy 1521 -----GGCAGGCCAGACAGAGAGCTGACAGAGCTGTTGACCGAG 1560
Db 728 AACGTTCCAGGGAGTCTGTTTGAAGCTTCCAGCGGAGAGCTGACAGTCTTCTGGAAG 787
Qy 1561 TACGGTCTTCTGAAGAGCTTCTGCACTCACAGCGCTTCTGTTTGGACTCTG 1620
Db 788 TATGGCAATTCGAAAGGGGTATTTGATTTTGTGTACGTTTCTCTTTGGGCTGGTA 847
Qy 1621 AACGAGGAGACAGGAGCCACTCTGGAAGAGTCTCTGCTGGAAGGTCTCGCCGCACATC 1680
Db 848 AACGAGGAGAGCCTCTACTTGGAGAAGAAATTAAGTTGCAAGATCTCTCAGCAATC 907
Qy 1681 AAGATGAGACCTGTGAGTGATCAAGAGAAAGCTCAGAGGAGGCTCCACCTGCGAG 1740
Db 908 AGGCTGGAGCTGCTGAATGGAATGAAGTGAAGCCAAAGCTAAAGAGCTGCAGATCCAG 967
Qy 1741 CAGGGCTCTTGGAGTCTTCTCAGCTGCTGTGACAGATCCAGGAGGAGGAGTTCATCCAG 1800
Db 968 CCCAGCAGCTGGAATGTTCTACTGTTTACGAGATGCAAGGAGGAGGAGTTCGTGCA 1027
Qy 1801 CAGGCCCTGAGCCACTCCAGAGTATGCTGTGCAACATTTGCTCCAGAGATGAGCAC 1860
Db 1028 AGGCCATGGAATTTCCCAAGATTGAG--ATCAATCTCTCCACAGAAATGGAACAC 1084
Qy 1861 ATGCTCTCTGCTGTCTGAGCGCTGAGAGGCGCCAGGCTGCTGCACTTGTATGGC 1920
Db 1085 ATGGTTTCTTCTTTGCAATGCAATGTCATCGGGTGGAGTCACTGTCCCTGGGGTTT 1144
Qy 1921 GCCACTACAGCGGGAGGAGAG----- 1946
Db 1145 CTCATATACATGCCCAG 1204
Qy 1947 CCGCGCAGAGTGTCTCCGAGGAGCGCACACGCTGTGTGAGCTTCAGACACAGAGAGGAC 2006
Db 1205 GTGAGTGTGCTCTCCAGAGCTCTCTCATGCTGCTGCTGCTCAT-GGATTTGGTGAACAG 1263
Qy 2007 CGTTCGTGAGAGCGCTACAGTGAACATCTGCGAGCGGCGCTGTGACCAATTCCAAACCT 2066
Db 1264 CCACCTACTTCCAGTCTTTTGGCGGGGCTCTTTTCACTTCTGAGCACAGCAGCAGAGTCT 1323
Qy 2067 GATAGAGCTGTCTGTACCGAATGCTTGGGAGCGCGGGGTGAGCTGCTCTGCA 2126
Db 1324 AACTGAATGGACCTCAGTGAACATTTCTCTGAGGAGCCAGGATGAGAGTGTGTGTGA 1383
Qy 2127 AGGATCAGACACCCCAACTTCAGAACTTCAGAACTGAGGCTGGAAGAGGTGCCGATCTC 2186
Db 1384 AACGCTCCAGCATCTGGCTGTAACTTCGGAGATTGTGTTGGGCGCTGTGGCTCTC 1443
Qy 2187 CAGCTCAGCTCGGAGGAGCTCTCTGAGCTCTCATAGCCAAATGAATTTGAAGAAGAT 2246
Db 1444 GCATGAGTGTCTGTCGACATCTCCTTGTGCTCAGAGCAACAGAGCTGTTGGAGCT 1503
Qy 2247 GGATCTCAGTGCACGGGTTGGATTCCAGGATGATGCTGCTTTGGAGGGCTGG 2306
Db 1504 GGACCTGAGTGACAAACGCGCTCGGTGACTTCGGAATCAGACTTCTGTGTGGGAGTAA 1563
Qy 2307 GCATCCCGAGTGCAGGCTGCAGATGATTCAAGTTGAGGAAGTGTACAGCTGGAGTCCGGGCG 2366

Db 1564 GCACCTGTGTGCAATCTGAAGAAGCTCTGGTTGGTCAGCTGCTCCTCACATCAGCAATG 1623
Qy 2367 TTGTCAAGAGATGCTTCTGTGCTCGGACCAACCCACATCTGGTTGAGTTGGACCTGAC 2426
Db 1624 TTGTCAAGGATCTTGCATCAGTATTGAGCAACAGGCAATTCCTCTGACAGACTCTATGTGGG 1683
Qy 2427 AGGAATGTCACTGAGGAGATTTGGGCTGTGAGGTACTATGCCAGGAGCTGAGGCAACCCAGT 2486
Db 1684 GGAGATGCTTGGAGACTCAGGAGTCGCAATTTATGTGAAAAGCAAGAAATCCACA 1743
Qy 2487 CTGAGACTACGGACTTTTGGCTGAAAGATGCGGCTCACTGCTGCTGCTGTGAGCA 2546
Db 1744 GTGTAACTGCAAGAACTGGGTTGGTGAATTTCTGGGCTTACGTCACTCTGTTGTTCCAG 1803
Qy 2547 GCTGGCTCAACTCTCAGTGTGAACAGAGCTCAGAGAGCTGAGACCTGAGCCTGAATGA 2606
Db 1804 TTTGCTCTCGTACTCAGCACTAATCAGAAATCTCAGCACTTTACCTGCGAGGCAACAC 1863
Qy 2607 GCTGGGAGCCTCGGGTGTCTGCTGTGTGAGGGCTCAGGCATCCAGTGTCAAGCT 2666
Db 1864 TCTCGAGACAAGGGATCAAACTACTCTGTGAGGAGTCTTGCACCCGACTGCAAGCT 1923
Qy 2667 CCAGACCTCGGGTTGGGCTGCTGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTTCTGT 2726
Db 1924 TCAGGTGTGGAATTAGACAACCTGCAACCTCACCTCACTGCTGCTGGGATCTTTCCAC 1983
Qy 2727 GGTCTCCAGGCCAACACACACTCCGGAGCTCGAGCTTGAGTTTCAACGACCTGGGAGA 2786
Db 1984 ACTTCTGACCTCCAGCAGAGCTGCGAAAGCTGAGCCTGGGCAACAATGACCTGGGGGA 2043
Qy 2787 CTGGGGCTGTGTGTGCTGCTGAGGGGCTGCAACATCCCGCTGCGAGCTCCAGAAAT 2846
Db 2044 CTTGGGGTCTATGATTTCTGTGAAGTCTGAACAGCAGAGCTGCCCTCTGCGAGACT 2103
Qy 2847 GTGGCTG 2853
Db 2104 GGGGTG 2110

RESULT 14
ID ABZ73494
XX ID ABZ73494 standard; cDNA; 2847 BP.
XX AC ABZ73494;
XX XX
XX DT 12-MAY-2003 (first entry)
XX DE Secreted protein-encoding gene 214 cDNA clone HNFY57, SEQ ID NO:224.
XX KW Human; secreted protein; cancer; tumour; hyperproliferative disorder;
XX KW autoimmune disorder; inflammation; angiogenic diseases; AIDS;
XX KW acquired immunodeficiency syndrome; hepatitis; anaemia; wound healing;
XX KW drug screening; chromosome identification; chromosome mapping;
XX KW cytosolic; gene therapy; antiinflammatory; immunomodulator; anti-HIV;
XX KW antianaemic; vulnery; chromosome 1q44; gene; ss.
XX OS Homo sapiens.
XX PN WO200277013-A2.
XX PD 03-OCT-2002.
XX PF 26-MAR-2002; 2002WO-US009370.
XX PR 27-MAR-2001; 2001US-0278650P.
XX PR 12-SEP-2001; 2001US-00950082.
XX PR 12-SEP-2001; 2001US-00950083.
XX PA (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Ruben SM;
XX WPI; 2003-040578/03.

P-PSDB; ABR01160.

New human secreted proteins and nucleic acids, useful for detecting or treating cancer or other hyperproliferative disorders, autoimmune disorders, inflammatory disorders, HIV disease, hepatitis or anemia.

Claim 21; Page 1274-1275; 247app; English.

ABZ73281-ABZ73697 represent cDNAs corresponding to 391 human secreted protein genes, and ABP00947-ABP01363 represent the proteins they encode. ABZ73698-ABZ74687 represent human secreted protein genomic fragments. The invention also encompasses antibodies specific for the secreted proteins, the use of the secreted proteins in drug screening and recombinant vectors and host cells comprising a nucleic acid of the invention. The secreted proteins are thought to be involved in biological activities associated with cellular signalling, cellular differentiation, cell migration, prohormone activation and neurotransmitter activity. The secreted proteins, nucleic acids encoding them, antibodies or antibody fragments specific for the secreted proteins, and modulators of protein activity are useful for diagnosing or treating cancers or other hyperproliferative disorders. Additionally, the secreted proteins and their nucleic acids may also be used in the treatment of autoimmune disorders, inflammatory disorders, diseases involving angiogenesis, AIDS (acquired immunodeficiency syndrome), hepatitis, anaemia, and to promote wound healing. Nucleic acids of the invention may be used for chromosome identification, chromosome mapping, in gene therapy, for identifying individuals from minute biological samples, as hybridisation probes, and as molecular weight markers. The present sequence represents a human secreted protein-encoding cDNA clone of the invention

Sequence 2847 BP; 706 A; 681 C; 719 G; 740 T; 0 U; 1 Other;

Query Match 19.1%; Score 594.6; DB 7; Length 2847;
Best Local Similarity 57.1%; Pred. No. 4.9e-137;
Matches 1203; Conservative 0; Mismatches 839; Indels 65; Gaps 4;

808 GATGAGCTCAGGCTCTTTCCAGCATCTCCAGGACCTCCAGGACCTGGTGCCTCTCTGGGAGGAG 867
8 GATGAGCTCAGGCTCTTTGACGACACATGAGGACCTCTGCACTGATGAGGAGGAG 67
868 AAACGGCCACGAGGAGTGTCTTTAAAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTA 927
68 GCCGAGGGGGAGACATCTCTGAGCAGCTCATCAGAAAGAGCTGCTCCGAGGCC 127
928 TCTTTGCTCATACACAGCGGCCAGCTTTGAGAGGCTCCACGCTCTGCTGGAGCAC 987
128 TCTCTGCTCATACACAGGACCTGTGGCCCTGAGAGAACTGCGACACTTGTCTGGACCAT 187
988 CCAGGAGCTGTGAGATCTCTGGGCTTCTGAGGACGAGGAGGAGGAGGAGGAGGAGGAG 1047
188 CCTGGCATGTGAGATCTCTGGGTTTCTCGAGGCGCAAGAGGAGGAGGAGGAGGAGGAGG 247
1048 TATTTCCACAATCAGAGGAGGCGGCGCAAGTCTTCAATACGTGAGGAGCAACAGGACCT 1107
248 TACTTCTGATGAGGCGCCCAAGCCAGGCGACCTTCACTGATTCAGGAGACAGGCTC 307
1108 CTCTTACCATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1167
308 CTCTTACCATGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 367
1168 CAGCTGGAGGCTGGGGCTGTTGAGACAGACCTTCCAGGACCAACCACTGAGTGTACATG 1227
368 CAGATGGAGTGGGAGAGGCTTGGCCAGACATTCAGAGACCAACCACTGAGTGTACATG 427
1228 CTCTACCTGTGAGTCTGATGCAACCAAGCCGCGGGGCGCCGCGGCTTCCAGCGCCCAACCC 1287
428 TTCTTCTTCCAGTGTGCTGAGCCCGCGGGAGGAGCCAGGAGCAAGGCTCTGGCC 487
1288 AACCAAGAGGCTGTGCTCTTGGCGGAGATGGGCTCTGGAATCAGAAAAATCTTATTT 1347
488 CACCTCTGGGGGCTCTGCTCTTTGGGTCAGATGGAATCTGGAACCAAGAAAAATCTCTGTT 547
1348 GAGGAGCAGGACCTCCGGAAGACGCGCTAGAGCGGGAAGACGCTCTCTGCTTCTCTCAAC 1407

548 GAGGAGTCCGACCTCAGGAATCATGGACTGCAGAAGCGGATGTGTCTCTCTGCTGAGG 607
1408 ATGAACATCTTCCAGAAGGACATCACTGTGTGAGAGGTACTACAGCTTTCATCCACTTCAGT 1467
608 ATGAACCTGTTCMAAAGGAGTGGACTGCGAAGTTCACAGCTTTCATCCATGACT 667
1468 TTCAGGAATCTTTGAGCTATGTACTATATCTCTGAGAGGCGGAGGCGG----- 1520
668 TTCAGGAGTCTTTTGGCGCATGTACTACTCTGTGGAAGGAAAGGAAGGAGGACG 727
1521 -----GGCAGGCCACAGACCGAGCTTCCAGCGGAGCGTGACAGCTCTCTGGAAGAC 787
728 AACGTTCCAGGAGTCTGTTGAAGCTTCCAGCGGAGCGTGACAGCTCTCTGGAAGAC 787
1561 TAGCGGTTTCTGAAGGAGCTCTCTGCACTACAGCGCTTCCGTTCTGTTTGGACTCTCTG 1620
788 TATGGCAATTCMAAAGGAGTATTTGATTTTGTGTAGCTTTCTCTCTTTGGCTGGTA 847
1621 AACGAGGAGACCGAGGACCACTGGAGAAAGTCTCTGTGGAGGCTTCGCGCACATC 1680
848 AACGAGGAGGAGCTCTCTACTTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATC 907
1681 AAGATGAGCTGTGAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCACCTCGAG 1740
908 AGGCTGAGCTGTGAAATGATTTGAAGTGAAGCAAGCTTAAAGCTTAAAGCTCGAGATCCAG 967
1741 CAGGGCTCTCTGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
968 CCAGCGAGCTGGAATTTCTACTGTTGTAGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 1027
1801 CAGGCTCTGAGCCACTTCCAGGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
1028 AGGCGCATGAGCTATTTCCCAAGATTGAG--ATCAATCTCTCCACCAAGATGAGCAC 1084
1861 ATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
1085 ATGCTTCT 1144
1921 GCCACCTTACAGCGGAGCGGGAAGA----- 1946
1145 CTCTATAACATGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1204
1947 CCAGCGAGGCTGCTCCGAGGAGCGACACGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2006
1205 GTGAGTGTGCTCTCCCAAGCT 1263
2007 CGTCTGCTGAGCGCTTACAGTGAACATCTGCGAGCGGCGCTGTGCTGCTGCTGCTGCTGCTGCT 2066
1264 CCACCTTCT 1323
2067 GATAGAGCTGCT 2126
1324 AACTGATTTGAGCTCAGTGACAAATCTCTGCGGGGAGCCAGGAGTGAAGTGTGTGTGA 1383
2127 AGGACTCAGACACCCCACTGCAAACTTCAAGACCTGAGGCTGAAGAGTGCCTGCTGCTGCTGCT 2186
1384 AACGCTCCAGCATCTCTGCTGTAACATTCGAGATTTGTTGGGCGCTGTGCGCTCTC 1443
2187 CAGCTCAGCTGCGAGGAGCTCTCTGAGCTCTCATAGCCAAATGAAGATTTGACAGGAT 2246
1444 GCATGAGTGTCTCTGAGATCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCT 1503
2247 GATCTCAGTGGCAAGCGGCTTGGATTCCAGGATGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2306
1504 GGACCTGAGTGACAAAGCGCTCTGCTGCTCTGCTCTCTGCTCTCTGCTCTCTGCTCTCTGCT 1563
2307 GCATCCCGAGTGCAGCTCAGATGATTCAGTTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2366
1564 GCACCTGTTGTGCAATCTGGAAGAGCTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1623
2367 TTCTCAGGAGATGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2426

Db 728 AACGTTCCAGGAGCTGTTTGAAGCTTCCAGCGACGTGACAGTCTTCTGGAAGAAC 787
Qy 1561 TACCGTGTCTGAAAGAGCTTCTCGGCACTACACAGCGCTTCTGTTTGGACTCTCTG 1620
Db 788 TATGGCAATTCGAAAGGGTATTTGATTTTGTGTTAGTCTTCTTCTTGGCTGGTA 847
Qy 1621 AACGAGAGACGAGGACCTGAGGAGAGTCTCTGTTGAAAGTCTCGCCGACATC 1680
Db 848 AACGAGAGAGGACCTCTACTTGGAGAGAAATTAAGTTGCAAGATCTCTCAGCAATC 907
Qy 1681 AAGATGACCTGTGTGAGTGGATCCAAAGCAAGCTCAGAGCGAGCTCCACCTCGAG 1740
Db 908 AGGCTGAGGCTGTGAAATGGAATGGAAGTGAAGCCAAAGCTAAAGAGCTGCAATCCAG 967
Qy 1741 CAGGGCTTCTTGGAGTCTTTCAGCTGTGTGTACAGATCCAGAGGAGGAGTCTTATCCAG 1800
Db 968 CCAGCCAGCTGGAATGTTCTACTGTGTGTACAGATCCAGAGGAGGAGTCTTGTGCA 1027
Qy 1801 CAGCCCTGAGCCACTTCCAGGATGATCGTGTGAGCAACATTTGCTCCAGATGGAGAC 1860
Db 1028 AGGCGCATGGACTATTTCCCAAGATTGAG---ATCAATCTCTCCACAGAAATGGACAC 1084
Qy 1861 ATGCTCTCTCTGTTCTGCTGAGCGCTGAGGAGCGCCAGGCTGCTGCACTTGTATGGC 1920
Db 1085 ATGGTTTCTCTCTTTCATGAGAACTGTCATCGGGTGGAGTCACTGTCCCTGGGGTTT 1144
Qy 1921 GCCACCTACAGCGCGGAGGGAAG----- 1946
Db 1145 CTCATRAATGCCAAGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1204
Qy 1947 CCGGCGAGGTGCTCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2006
Db 1205 GTGCGAGTGTCTCTCCCAAGCTCTCTCATGCTGCTCTTCTCAT-GGAATTTGGTGAACAG 1263
Qy 2007 CGTTCCTGAGCGCTTACAGTGAACATCTGGCAGCGCCCTGTCACCAATCCAAACCT 2066
Db 1264 CCACCTCACTTCAGTTTTCGCGGGGCTCTTTTCAGTTCTGAGCACCAGCCAGAGTCT 1323
Qy 2067 GATAGAGCTGCTCTGTACGAAATGCCCTGGGAGCGCGGGGGTGAAGTGTCTCTCA 2126
Db 1324 AACTGAATTGAGCTCAGTGACAATTTCTTGGGGGAGCCAGGAGTGAAGTGTGTGA 1383
Qy 2127 AGGACTCAGACACCCCACTGCAAACTTCAAGCTGAGGCTGAAGAGTGGCGCATCTC 2186
Db 1384 AACGCTCCAGCATCTCGCTGTAACTTCGAGATTTGTGTTGGGGGCTGTGGCTCTC 1443
Qy 2187 CAGCTCAGCTGAGGAGCTCTCTGAGCTCTCATAGCCAAATAAGAAATTTGACAGGAT 2246
Db 1444 GCATGAGTGTCTTGCATCTCTTGGTCTCAGCAGCAACGAGAGCTGTGGAGCT 1503
Qy 2247 GATCTCAGTGGCAACGGCGTTGGAATCCAGGAGTATGCTGTTTGGAGGGCTTGGC 2306
Db 1504 GGACCTGAGTGACAAGCCCTCGTGACTTCGGAATCAGACTTCTGTGTGGGACTGAA 1563
Qy 2307 GCATCCCGAGTGAAGCTGAGATGATGATGTTGAGGAATGTCAGTGGAGTCCGGGGC 2366
Db 1564 GCACCTGTTGTGCAATCTGAAGAGCTCTGTGTTGGTCAAGCTGTGCTCATCAGCATG 1623
Qy 2367 TTGTCAGGAGATGGCTTCTGTCTCGGACCAACCCACATCTGGTTGAGTTGGACCTGAC 2426
Db 1624 TTGTCAGGATCTTGCATCAGTATGAGCAGCAGCATTCCTCAGCAGACTCTATGTTGG 1683
Qy 2427 AGGAATGCACTGAGGAGATTTGGCCCTGAGGTTACTATGCCAGGAGTGAAGGCCACCGT 2486
Db 1684 GGAGAAATGCCCTTGGGAGACTCAGGAGTCCGAAATTTTATGMAAAAGCCAGAAATCCACA 1743
Qy 2487 CTGCACTACGAGCTTTGTGGCTGAAGATCTGCCGCTCAGTCTGCTGCTGTGAGCA 2546
Db 1744 GTGTAACCTGCAGAAACTGGGGTGTGGAATTTCTGGCTTTACGTCAGTCTGTGTTCCAGC 1803
Qy 2547 GCTGGCCTCAACTCTCAGTGTGAACACAGAGCTGAGAGAGCTGGACCTGAGCCTGAATGA 2606

Db 1804 TTTGTCTCTCGGTACTCAGCACTAATCAGAAATCTCACGCACTTTTACCTGCGAGGCAACAC 1863
Qy 2607 GCTGGGGAGCTCTGGGGTGTCTGTGTGTGAGGGGCTCAGGCAATCCCACTGCAAGCT 2666
Db 1864 TCTCGAGAGACAGGGGATCAAACTACTCTGTGAGGAGTCTTGCACCCGACTGCAAGCT 1923
Qy 2667 CCAGACCTCTGGGTGGGCACTGTGCCGGTGGGCTCTGCGGCTGTGAGGGTCTTTCTGT 2726
Db 1924 TCAGGTGTTGGAAATTAGACAACTGCAACCTCACGTCACACTGCTGTCTGGGATCTTCCAC 1983
Qy 2727 GGTGCTCCAGGCCAACCAACCTCCGGGAGCTGGAATTTGAGTTTCAACGACCTGGGAGA 2786
Db 1984 ACTTCTGACTCCAGCCAGAGCTTCCGAAAGCTGAGCTTGGGCAACAAATGACCTTGGGCGA 2043
Qy 2787 CTGGGGCTCTGTGTTGCTGTGCTGAGGGGCTGCAACATCCCGCTTGCAGACTTCCAGAACT 2846
Db 2044 CCTGGGGTCAATGATGTTCTGTGAGTGTGAAACAGCAGAGCTGCTCTCTGCAAGACCT 2103
Qy 2847 GTGGCTG 2853
Db 2104 GGGGTTG 2110

Search completed: August 8, 2004, 01:41:11
Job time : 1147 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 8, 2004, 07:08:21 ; Search time 11827 Seconds
(without alignments)
11390.052 Million cell updates/sec

Title: US-10-781-294-23

Perfect score: 3108

Sequence: 1 atgtctagcaacgcaggcag.....atttgacattggctgctga 3108

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pin.*
35: em.htg.rod.*
36: em.htg.rod.*
37: em.htg.vrt.*
38: em.sv.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3108	100.0	3108	6	AX417214 Sequence
2	3057	98.4	3300	6	AX459869 Sequence
3	2820	90.7	3731	9	AY116204 Homo sapi
4	2577	82.9	3507	9	BC028069 Homo sapi
5	2562	82.4	3563	9	AY116205 Homo sapi
6	2475	79.6	3827	9	AY095146 Homo sapi
7	2475	79.6	3827	9	AY154467 Homo sapi
8	2223	71.5	3395	9	AY116206 Homo sapi
9	2220	71.4	3221	9	AY116207 Homo sapi
10	1995	64.2	3466	6	AX833594 Sequence
11	1995	64.2	3466	9	AC095460 Homo sapi
12	1703	54.8	147330	9	AC008753 Homo sapi
13	1356	43.6	2158	6	AX575503 Sequence
14	768	24.7	4931	6	AX684291 Sequence
15	515	16.6	2494	9	AF231021 Homo sapi
16	143	4.6	308	6	AX684295 Sequence
17	69	2.2	80	6	AX684299 Sequence
18	28	0.9	2159	10	AY364010 Mus muscu
19	28	0.9	218270	2	AC079499 Mus muscu
20	28	0.9	223344	2	AC106193 Rattus no
21	26	0.8	39	6	AX684309 Sequence
22	25	0.8	37	6	AX684310 Sequence
23	25	0.8	37	6	AX684312 Sequence
24	24	0.8	38	6	AX684311 Sequence
25	24	0.8	110000	2	AC146908 Oryza sat
26	24	0.8	135406	8	AC136843 Oryza sat
27	24	0.8	144077	2	AC135398 Oryza sat
28	24	0.8	179591	2	AP004805 Oryza sat
29	24	0.8	252463	2	AC139169 Oryza sat
30	23	0.7	227555	10	AC108416 Mus muscu
31	23	0.7	231864	2	AC097246 Rattus no
32	23	0.7	240461	2	AC111669 Rattus no
33	23	0.7	249541	2	AC095672 Rattus no
34	23	0.7	249774	2	AC126196 Rattus no
35	22	0.7	114098	9	AL162594 Human DNA
36	22	0.7	156094	9	AC090198 Homo sapi
37	22	0.7	183778	2	AC011004 Homo sapi
38	22	0.7	184548	2	AC023551 Homo sapi
39	22	0.7	193882	9	AF313041 Homo sapi
40	22	0.7	199437	2	AC091575 Homo sapi
41	22	0.7	207069	9	AC090922 Homo sapi
42	22	0.7	207190	9	AF295017 Homo sapi
43	22	0.7	209296	5	AL954150 Zebrafish
44	22	0.7	211018	5	BX004964 Zebrafish
45	22	0.7	218270	2	AC079499 Mus muscu

ALIGNMENTS

RESULT 1
AX417214
LOCUS
Sequence 23 from Patent WO0226780.
DEFINITION
AX417214
ACCESSION
AX417214.1 GI:21449784
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Reed, J.C., Godzik, A., Chu, Z.L., Pawlowski, K., Fiorentino, L.,
Ariza, M.E. and Stehlik, C.
Paad domain-containing polypeptides, encoding nucleic acids, and

methods of use		
JOURNAL	Patent: WO 0226780-A 23 04-APR-2002;	
FEATURES	The Burnham Institute (US)	
source	Location/Qualifiers	
	1..3108	/organism="Homo sapiens"
		/mol_type="unassigned DNA"
		/db_xref="taxon:9606"
CDS	1..3108	/note="unnamed protein product"
	/codon_start=1	
	/protein_id="CAD35284.1"	
	/db_xref="GI:21449785"	
	/db_xref="REMBL:CAD35284"	
	/translation="MLRTAGRDGLRLSLYLEELAEVLEKFKFLYLGTATLGLRGKPIP WGSMEKAGPLEMAQLLI THFGPEEARWRLALSTFERINRKNOLWERSQEDLVRDQETV RDYVRKFLMEDNRNARLIGECVNLSHRYTRLLLLYKESNPMQVQOQLDTRGHRATV GHQASPIKETLFEPEDEERPEPRIVMQGAAGIGKSLAHKVMMDWADGKLFQGRFD YLFIXNKQMSQATECSMDLIIFSCWPEPSAPLOELIRVERLLFIIDGFDELKPSF HDPQGNCLWEEKRPTLELNLIRKLLPELSLLITRPTALEKLRLHLEHPRHVE ILGFEAEKEYFYKYNARQAGVNYVRDNEPLFTWCFVPLVVCVCTCLOQOLE GGGLRQTSRTTAVWLYLISLMQPKGAPRIQPPNORGLCSLAADGLWNOKILFE EODLRKHGLDGEDVSALNNIIFQDINCERYYSFIHLSQEPFAANYIILDEGEGA GPDQVRLITETAPERSFLATSRFLFGLNSETSHLEKSLCWKVSPIKMDLIQ WIQKASDGLTQOQSLFEPSCLYEIOEEFIQOALSHFQVIYVSNIAKMEHWS FCLRCRSQVILHYGATYSADGEDRCRCSAGATHLLVQLRPERITVLLDAYSEHLAA LCTNPNIELSLYRNALGSRGVLLKLCQGLRHPNCKLQNLRLKRCRISSACEIDLAA IANKNLTRMDSGNVGPFGPMMLLCEGLRHPQCRLOMLOLRKCOLESQAQOEMASVLG TNPHVELDLTGNALEDGLRLCQGLRHPVCRILTWLKICRLTAACDELASTLSV NQSLRELDLSNELGDGLVLLCGLRHPCKLQTLRIGICRLGSAACBSLSVLIQAN HNLRDLDSFDNLGDWGLMLAEGLOHPACRLQKMLDSCGLTAKACENLYFTLIGINO TLTDLXLTNNALGDTGVRLLCKRLSHFPCCKLRVLWLFQMDLNKMTSHRLAALRVTKPY LDIGC"	
ORIGIN		
Query Match	100.0.%; Score 3108; DB 6; Length 3108;	
Best Local Similarity	100.0.%; Pred. No. 0;	
Matches 3108; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	ATGCTACGAACCGCAGGAGGAGCGCCCTCTGTCGCCTGTCCACCTACTTGGAGAATC 60
DB	1	ATGCTACGAACCGCAGGAGGAGCGCCCTCTGTGCGCTGTCCACCTACTTGGAGAATC 60
QY	61	GAGGCTGTGGAATCGAAGAAGTTCAAGTTATACCTGGGGACCGCAGACAGAGCTGGAGAA 120
DB	61	GAGGCTGTGGAATCGAAGAAGTTCAAGTTATACCTGGGGACCGCAGACAGAGCTGGAGAA 120
QY	121	GGCAAGATCCCTGGGGAAGCATGAGNAGCCCGTTCCTCTGGAAATGGCCCAAGCTGCTC 180
DB	121	GGCAAGATCCCTGGGGAAGCATGAGNAGCCCGTTCCTCTGGAAATGGCCCAAGCTGCTC 180
QY	181	ATCACCACCTTGGGGCCAGAGGAGGCTTGGAGGTGGCTCTCAGCACCTTTGAGGGATA 240
DB	181	ATCACCACCTTGGGGCCAGAGGAGGCTTGGAGGTGGCTCTCAGCACCTTTGAGGGATA 240
QY	241	AACAGGAAGGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAGGGATCCCGAGAA 300
DB	241	AACAGGAAGGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAGGGATCCCGAGAA 300
QY	301	ACCTACAGGGAATGTGTCGCCAGGAATAATTCGGGCTCATGGAGACCGCAATGCGCGCTA 360
DB	301	ACCTACAGGGAATGTGTCGCCAGGAATAATTCGGGCTCATGGAGACCGCAATGCGCGCTA 360
QY	361	GGGGAATGTGTCAACTCAGCCACCGGTACACCCGGGCTCTGTGCTGGTGAAGGACCTCA 420
DB	361	GGGGAATGTGTCAACTCAGCCACCGGTACACCCGGGCTCTGTGCTGGTGAAGGACCTCA 420
QY	421	AACCCCATGAGTCCAGCAGACGCTTCTGTGACACAGCCGGGAGACACGCGAGACCGTG 480
DB	421	AACCCCATGAGTCCAGCAGACGCTTCTGTGACACAGCCGGGAGACACGCGAGACCGTG 480
QY	481	GGACACGAGCTAGCCCATCATAGATAGAGACCTCTTTGAGCCAGACGAGGAGCGGCC 540
DB		

481	GGACACGAGCTAGCCCATCATAGATAGAGACCTCTTTTGGCCAGACGAGGAGCGCCCC	541
QY	GAGCCACCGCGCACCGTGGTGCATGCAAGCGCGGAGGATAGGCAAGTCCATCTGTGGCA	600
DB	GAGCCACCGCGCACCGTGGTGCATGCAAGCGCGGAGGATAGGCAAGTCCATCTGTGGCA	600
QY	CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAAGGCAGATTTGATTATCTC	660
DB	CACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAAGGCAGATTTGATTATCTC	660
QY	TTCTACATCAACTGCAAGGAGATGAACCAAGAGTGCACGGAATGCAAGACCTC	720
DB	TTCTACATCAACTGCAAGGAGATGAACCAAGAGTGCACGGAATGCAAGACCTC	720
QY	ATCTTCAGCTGCTGGCTGAGCCGCGCTCTCCAGAGCTCATCCGAGTCCCGAG	780
DB	ATCTTCAGCTGCTGGCTGAGCCGCGCTCTCTCCAGAGCTCATCCGAGTTCCTCCAG	780
QY	CGCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCCTTCTTCCACGATCCTCAG	840
DB	CGCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCCTTCTTCCACGATCCTCAG	840
QY	GGACCTCTGCTGCTGCTGGGAGGAGAAACGGCCACCGAGCTGCTTCTTAACAGCTT	900
DB	GGACCTCTGCTGCTGCTGGGAGGAGAAACGGCCACCGAGCTGCTTCTTAACAGCTT	900
QY	ATTTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCACACGGCCACGGCTT	960
DB	ATTTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACCACACGGCCACGGCTT	960
QY	GAGAAGCTCCACCGCTGCTGGAGCACCCACGAGCATGTGGAGATCCTGGGCTTCTCTGAG	1021
DB	GAGAAGCTCCACCGCTGCTGGAGCACCCACCGAGCATGTGGAGATCCTGGGCTTCTCTGAG	1021
QY	GCAGAAAGGAAGAAATCTTCTTCAAGTATTTCCAAATGCAAGACGAGCGGCGCAAGTC	1081
DB	GCAGAAAGGAAGAAATCTTCTTCAAGTATTTCCAAATGCAAGACGAGCGGCGCAAGTC	1081
QY	TTCAATTAAGTGTGAGGAGCAACGAGCTCTCTTCAACATGCTGCTTCCCTCCCTGCTGTC	1141
DB	TTCAATTAAGTGTGAGGAGCAACGAGCTCTCTTCAACATGCTGCTTCCCTCCCTGCTGTC	1141
QY	TGGGTGCTGTGTACCTTGCCTCCAGACAGCTGGAGGGTGGGGGGCTGTTGAGACAGAGC	1201
DB	TGGGTGCTGTGTACCTTGCCTCCAGACAGCTGGAGGGTGGGGGGCTGTTGAGACAGAGC	1201
QY	TCCAGAACCACTGCTGAGTGTACATGCTTCTACCTGCTGAGTCTGATGCAACCCCAAGCG	1261
DB	TCCAGAACCACTGCTGAGTGTACATGCTTCTACCTGCTGAGTCTGATGCAACCCCAAGCG	1261
QY	GGGGCCCCGGCTCCAGCCCCCAACCCACAGAGAGGGTGTGCTCTTGGGGGAGAT	1321
DB	GGGGCCCCGGCTCCAGCCCCCAACCCACAGAGAGGGTGTGCTCTTGGGGGAGAT	1321
QY	GGGCTCTGGAATCAGAAAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGAC	1381
DB	GGGCTCTGGAATCAGAAAAATCCTATTGAGGAGCAGGACCTCCGGAAGCAGCGGCTAGAC	1381
QY	GGGGAAGAAGCTCTGCTCCCTTCAACATGAACATCTTCCAGAGGACATCAACTGTGAG	1441
DB	GGGGAAGAAGCTCTGCTCCCTTCAACATGAACATCTTCCAGAGGACATCAACTGTGAG	1441
QY	AGGTACTACAGCTTCACTTGAATTTCCAGGAATCTTTTGGAGCTATGTACTATATC	1501
DB	AGGTACTACAGCTTCACTTGAATTTCCAGGAATCTTTTGGAGCTATGTACTATATC	1501
QY	CTGGACGAGGGGAGGGCGGGCAGSCCCAGACACGAGGACGTGACCAAGCTGTTGACCCAG	1561
DB	CTGGACGAGGGGAGGGCGGGCAGSCCCAGACACGAGGACGTGACCAAGCTGTTGACCCAG	1561
QY	TACGCGTTTTTGAAGAGGAGCTTCTTGGCACTCAACAGCCGCTTCTGTTGGACTCTCT	1621
DB	TACGCGTTTTTGAAGAGGAGCTTCTTGGCACTCAACAGCCGCTTCTGTTGGACTCTCT	1621

Db	481	GGACACGAGGCTAGCCCCCATCAAGATAGAGACCTCTTTTGAGCCAGACGAGGAGCGCCCC 540
Qy	541	GAGCACCGCCACCGTGGTTCATCAAGCGCGCGCAGGATAGGCAAGTCCATCTCTGCA 600
Db	541	
Db	541	GAGCACCGCCACCGTGGTTCATCAAGCGCGCGCAGGATAGGCAAGTCCATCTCTGCA 600
Qy	601	CACAAGGTGATGCTGGACTGGCGGAGCGGGAAGCTCTTCCAAAGGAGATTTGATTATCTC 660
Db	601	
Db	601	CACAAGGTGATGCTGGACTGGCGGAGCGGGAAGCTCTTCCAAAGGAGATTTGATTATCTC 660
Qy	661	TTCTACATCAATCGAGGAGATGAAACAGAGTGCCACGGAATGAGCATGCAAGACCTC 720
Db	661	
Db	661	TTCTACATCAATCGAGGAGATGAAACAGAGTGCCACGGAATGAGCATGCAAGACCTC 720
Qy	721	ATCTTCAGCTGCTGGCCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCTCAG 780
Db	721	
Db	721	ATCTTCAGCTGCTGGCCTGAGCCAGCGCGCTCTCCAGGAGCTCATCCGAGTTCCTCAG 780
Qy	781	CGCCTCTCTTTTCATCTGACGCGCTTCGATGAGCTCAAGCCTTCTTCCAGATCTCTCAG 840
Db	781	
Db	781	CGCCTCTCTTTTCATCTGACGCGCTTCGATGAGCTCAAGCCTTCTTCCAGATCTCTCAG 840
Qy	841	GGACCTCTGGTCTCTGCTGGGAGAGAAACGGGCCACGAGAGCTGCTTCTTAAACAGCTTA 900
Db	841	
Db	841	GGACCTCTGGTCTCTGCTGGGAGAGAAACGGGCCACGAGAGCTGCTTCTTAAACAGCTTA 900
Qy	901	ATTGCGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGCGCCACCGCTTTG 960
Db	901	
Db	901	ATTGCGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGCGCCACCGCTTTG 960
Qy	961	GAGAAGTCTCAACCGTCTGCTGGAGCACCCACAGGATGTGGAGATCTTGGGCTTCTCTGAG 1020
Db	961	
Db	961	GAGAAGTCTCAACCGTCTGCTGGAGCACCCACAGGATGTGGAGATCTTGGGCTTCTCTGAG 1020
Qy	1021	GCAGAAAGAGGAATATCTTCAAGTATTTCCAAATGACAGAGAGCGGGCCCAAGTC 1080
Db	1021	
Db	1021	GCAGAAAGAGGAATATCTTCAAGTATTTCCAAATGACAGAGAGCGGGCCCAAGTC 1080
Qy	1081	TTCAATTTACGTGAGGAGAACAGAGCTCTTTCAACCATGTCTTCTGCTCCCTCGTGTGC 1140
Db	1081	
Db	1081	TTCAATTTACGTGAGGAGAACAGAGCTCTTTCAACCATGTCTTCTGCTCCCTCGTGTGC 1140
Qy	1141	TGGTGTGTGTACTCTCTCAGCAGAGCTGGAGGCTGGGGGCTGTTTGAGACAGACG 1200
Db	1141	
Db	1141	TGGTGTGTGTACTCTCTCAGCAGAGCTGGAGGCTGGGGGCTGTTTGAGACAGACG 1200
Qy	1201	TCCAGGACCACTGTCAGTGTATCTCTTACCTGTGAGTCTGATGCAACCAAGCCG 1260
Db	1201	
Db	1201	TCCAGGACCACTGTCAGTGTATCTCTTACCTGTGAGTCTGATGCAACCAAGCCG 1260
Qy	1261	GGGCCCCCGGCTTCCAGCCCCCAACCAAGAGAGGTTGTCTCTCTTGGCGGAGAT 1320
Db	1261	
Db	1261	GGGCCCCCGGCTTCCAGCCCCCAACCAAGAGAGGTTGTCTCTCTTGGCGGAGAT 1320
Qy	1321	GGGCTCTGGAATCAGAAAAATCTTATTTGAGAGCAGACCTCCGGAAGACGCGCTAGAC 1380
Db	1321	
Db	1321	GGGCTCTGGAATCAGAAAAATCTTATTTGAGAGCAGACCTCCGGAAGACGCGCTAGAC 1380
Qy	1381	GGGGAAGACCTCTGCTGCTTCTCAACATGAACATCTTCCAGAGGACATCACTGTGAG 1440
Db	1381	
Db	1381	GGGGAAGACCTCTGCTGCTTCTCAACATGAACATCTTCCAGAGGACATCACTGTGAG 1440
Qy	1441	AGGTACTTACAGCTTCATCTGAGTTTCCAGGAATCTTTTGAGCTATGTACTATATC 1500
Db	1441	
Db	1441	AGGTACTTACAGCTTCATCTGAGTTTCCAGGAATCTTTTGAGCTATGTACTATATC 1500
Qy	1501	CTGACGAGGGGAGGCGGGCAGGCCAGACAGACAGACGTCGACAGGCTTGTGACCGAG 1560
Db	1501	
Db	1501	CTGACGAGGGGAGGCGGGCAGGCCAGACAGACAGACGTCGACAGGCTTGTGACCGAG 1560
Qy	1561	TAGCGTTTCTGAAGAGGCTCTCTGCGACTCACAGCGCTTCTCTGTTGGACTCTCTG 1620
Db	1561	
Db	1561	TAGCGTTTCTGAAGAGGCTCTCTGCGACTCACAGCGCTTCTCTGTTGGACTCTCTG 1620

1621 AACGAGGAGACGAGGACCACTGAGAGAGTCTCTGCTGGAAGGTCTCGCGGCACATC 1680
1621 AACGAGGAGACGAGGACCACTGAGAGAGTCTCTGCTGGAAGGTCTCGCGGCACATC 1680
1681 AAGATGACCTGTTGAGTGGATCAAAAGCAAGCTCAGAGGACGCTCCACCTGCGAG 1740
1681 AAGATGACCTGTTGAGTGGATCAAAAGCAAGCTCAGAGGACGCTCCACCTGCGAG 1740
1741 CAGGCTCTCTGAGTCTCTCAGCTGCTGTGACGAGTCCAGAGGAGGAGTTATCCAG 1800
1741 CAGGCTCTCTGAGTCTCTCAGCTGCTGTGACGAGTCCAGAGGAGGAGTTATCCAG 1800
1801 CAGGCTCTGAGGACCTTCCAGCTGATCGTGGTCAGCAACATTTGCCCTCCAGATGAGCAC 1860
1801 CAGGCTCTGAGGACCTTCCAGCTGATCGTGGTCAGCAACATTTGCCCTCCAGATGAGCAC 1860
1861 ATGGTCTCTCTGTTCTGTAAGCGCTGACAGAGCGCCAGGTCGTGCACTTGTATGGC 1920
1861 ATGGTCTCTCTGTTCTGTAAGCGCTGACAGAGCGCCAGGTCGTGCACTTGTATGGC 1920
1921 GCCACTACAGCGGAGCGGGAAGACCGCGAGGTCTCCGAGGAGCGCACAGCTG 1980
1921 GCCACTACAGCGGAGCGGGAAGACCGCGAGGTCTCCGAGGAGCGCACAGCTG 1980
1981 TTGGTGCAGCTCAGACCAAGAGGACCGTCTCTGCTGGAGCGCTCAGTGAACTTGGA 2040
1981 TTGGTGCAGCTCAGACCAAGAGGACCGTCTCTGCTGGAGCGCTCAGTGAACTTGGA 2040
2041 GCGGCTCTGTGACCAATCCAACTGATAGAGTGTCTGTATCGGAAATGCCCTGGGC 2100
2041 GCGGCTCTGTGACCAATCCAACTGATAGAGTGTCTGTATCGGAAATGCCCTGGGC 2100
2101 AGCGGGGGGTGAAGCTGCTCTGTAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
2101 AGCGGGGGGTGAAGCTGCTCTGTAAGGACTCAGACACCCCACTGCAAACTTCAGAAC 2160
2161 CTGAGGCTGAAGAGGTGCGGACATCTCCAGCTCAGGAGTCTCTGCTGAGTCTC 2220
2161 CTGAGGCTGAAGAGGTGCGGACATCTCCAGCTCAGGAGTCTCTGCTGAGTCTC 2220
2221 ATAGCCAAATAGAAATTTGAAAGGATGATCTCAGTGGCAACGGGTTGGATTCACAGC 2280
2221 ATAGCCAAATAGAAATTTGAAAGGATGATCTCAGTGGCAACGGGTTGGATTCACAGC 2280
2281 ATGATGCTGCTTTGAGGGGCTGCGGATCCCACTGAGGAGTCTCAGTGTGAGTTCAGT 2340
2281 ATGATGCTGCTTTGAGGGGCTGCGGATCCCACTGAGGAGTCTCAGTGTGAGTTCAGT 2340
2341 AGGAAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2400
2341 AGGAAAGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2400
2401 CCACATCTGTTGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2460
2401 CCACATCTGTTGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2460
2461 CTATCCAGGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2520
2461 CTATCCAGGAGTGGAGTCCGGGGCTTGTGAGGAGTGGCTTCTGCTCGGACCAAC 2520
2521 CGCCTCACTGCTGCTGTGACAGGCTGGGCTTCAACTCTCAGTGTGAAACAGAGCTG 2580
2521 CGCCTCACTGCTGCTGTGACAGGCTGGGCTTCAACTCTCAGTGTGAAACAGAGCTG 2580
2581 AGAGAGTGGAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2640
2581 AGAGAGTGGAACTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2640
2641 GGCCTCAGGCATCCACGTCGAGCTCCAGACCTGCGGTTGGGCTATGCGGCTGGGC 2700
2641 GGCCTCAGGCATCCACGTCGAGCTCCAGACCTGCGGTTGGGCTATGCGGCTGGGC 2700

2701 TCTGCGGCTGTGAGGGTCTTTCTGCTGCTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
2701 TCTGCGGCTGTGAGGGTCTTTCTGCTGCTGCTCCAGGCAACCAACCTCCGGGAGCTG 2760
2761 GACTTGAGTTCACAGCACTGCGGAGTCTGCGGCTGCTGCTGCTGAGGAGCTGCA 2820
2761 GACTTGAGTTCACAGCACTGCGGAGTCTGCGGCTGCTGCTGCTGAGGAGCTGCA 2820
2821 CATCCCGCTGCAGACTCCAGAACTGTGGCTGGATAGCTGTGGCTCAGACCAAGGCT 2880
2821 CATCCCGCTGCAGACTCCAGAACTGTGGCTGGATAGCTGTGGCTCAGACCAAGGCT 2880
2881 TGTGAGAACTCTTACTTCAACCTGCGGATCAACAGAGCTTTCACCGACCTTTACCTGACC 2940
2881 TGTGAGAACTCTTACTTCAACCTGCGGATCAACAGAGCTTTCACCGACCTTTACCTGACC 2940
2941 AACACGCTCTAGGAGGACACAGGTGTCCGACTGCTTTTGAAGCGGTGAGCCATCTGCGC 3000
2941 AACACGCTCTAGGAGGACACAGGTGTCCGACTGCTTTTGAAGCGGTGAGCCATCTGCGC 3000
3001 TGCAAACTCCGAGTCTCTGCTTATTTGGGATGGACCTGAATAAATGACCCACAGTAGG 3060
3001 TGCAAACTCCGAGTCTCTGCTTATTTGGGATGGACCTGAATAAATGACCCACAGTAGG 3060
3061 TTGGCAGCGCTTCGAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108
3061 TTGGCAGCGCTTCGAGTAACAAACCTTTATTTGGACATTTGGCTGCTGA 3108

RESULT 2
AX459869
LOCUS AX459869 3300 bp DNA linear PAT 30-NOV-2002
DEFINITION Sequence 12 from Patent WO0240668.
ACCESSION AX459869
VERSION AX459869.1 GI:21725643
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Tschoep, J. and Martinon, F.
Proteins and dna sequences underlying these proteins used for
treating inflammations
Patent: WO 0240668-A 12 23-MAY-2002;
Apotech Research and Development Ltd. (CH)
Location/Qualifiers
1. .3300
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="NALP3/PY5.dna"

ORIGIN
Query Match 98.4%; Score 3057; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3057; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGTACGAAACGCGAGGAGGACCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
Db 1 ATGTACGAAACGCGAGGAGGACCGGCTCTGTGCGCTGTCCACCTACTTGGAGAACTC 60
QY 61 GAGGCTGTGGAACTCAAGAAGTTCAAGTTATACCTGGGAGCGCCACAGAGCTGGAGAA 120
Db 61 GAGGCTGTGGAACTCAAGAAGTTCAAGTTATACCTGGGAGCGCCACAGAGCTGGAGAA 120
QY 121 GGCAAGATCCCTCGGGAGGATGAGAGGCGGCTGCCCTGGAAATGSCCAGCTGCTC 180
Db 121 GGCAAGATCCCTCGGGAGGATGAGAGGCGGCTGCCCTGGAAATGSCCAGCTGCTC 180
QY 181 ATCACCCTACTTGGGCCAGAGAGGCTTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
Db 181 ATCACCCTACTTGGGCCAGAGAGGCTTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240

QY 241 AACGAGGAGGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAAGGATCCCCAGGAA 300
DB 241 AACGAGGAGGACCTGTGGGAGAGAGGACAGAGAGAGGACCTGGTGAAGGATCCCCAGGAA 300
QY 301 ACCTACAGGAGCTATGTCCGAGAGAAATTCGGGTCTATGGAAGACCGCAATGCGGCGCTA 360
DB 301 ACCTACAGGAGCTATGTCCGAGGAAATTCGGGTCTATGGAAGACCGCAATGCGGCGCTA 360
QY 361 GGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCGTGGTGAAGGACACTCA 420
DB 361 GGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCGTGGTGAAGGACACTCA 420
QY 421 AACCCCATGACAGGTCCAGCAGCAGCTTCTGACACAGCCCGGGGACACGCGAGGACCGTG 480
DB 421 AACCCCATGACAGGTCCAGCAGCAGCTTCTGACACAGCCCGGGGACACGCGAGGACCGTG 480
QY 481 GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAGGAGCGGCC 540
DB 481 GGACACAGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAGGAGCGGCC 540
QY 541 GAGCACCGCCACCGTGTATGCAAGGCGCGGAGGATAGGCAAGTCCATGCTGGCA 600
DB 541 GAGCACCGCCACCGTGTATGCAAGGCGCGGAGGATAGGCAAGTCCATGCTGGCA 600
QY 601 CACAAGGTGATGTGGACTGGGCGGAGCGGGAAGCTTTTCAAGGCAGATTGATTATCTC 660
DB 601 CACAAGGTGATGTGGACTGGGCGGAGCGGGAAGCTTTTCAAGGCAGATTGATTATCTC 660
QY 661 TTCTACATCAATCGAGGAGATGAAACAGAGTGGCAGCGAATGAGCATGCAAGCTC 720
DB 661 TTCTACATCAATCGAGGAGATGAAACAGAGTGGCAGCGAATGAGCATGCAAGCTC 720
QY 721 ATCTTTAGCTGCTGCTGAGCCAGCCAGCGCCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
DB 721 ATCTTTAGCTGCTGCTGAGCCAGCCAGCGCCTCTCCAGGAGCTCATCCGAGTTCCCGAG 780
QY 781 CGCTCTCTTTTCAATCATCGACGGCTTCGATGAGCTCAAGCTTTCTTCCAGATCCTCAG 840
DB 781 CGCTCTCTTTTCAATCATCGACGGCTTCGATGAGCTCAAGCTTTCTTCCAGATCCTCAG 840
QY 841 GGACCTGTGCTCTGCTGGGAGAGAAACGGCCACGAGCTGCTTTTAAAGCTTA 900
DB 841 GGACCTGTGCTCTGCTGGGAGAGAAACGGCCACGAGCTGCTTTTAAAGCTTA 900
QY 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCGCCAGCGCTTG 960
DB 901 ATTGGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGGCGCCAGCGCTTG 960
QY 961 GAGAAGCTCCACCTGCTGAGCACCCAGGCAATGGAAGATGTTGAGATCCTGGGCTTCTTGAG 1020
DB 961 GAGAAGCTCCACCTGCTGAGCACCCAGGCAATGGAAGATGTTGAGATCCTGGGCTTCTTGAG 1020
QY 1021 GCAGAAAGGAGAAATCTTCAAGTATTTTCAATGAGAGAGAGCGGCGGCAAGTC 1080
DB 1021 GCAGAAAGGAGAAATCTTCAAGTATTTTCAATGAGAGAGAGCGGCGGCAAGTC 1080
QY 1081 TTCAATTAAGTGAGGAGCAACAGGCTCTTTTCAATGAGTCTGCTCCCTCGTGTGC 1140
DB 1081 TTCAATTAAGTGAGGAGCAACAGGCTCTTTTCAATGAGTCTGCTCCCTCGTGTGC 1140
QY 1141 TGGTGTGTGTACTGCTCCAGCAGCAGCTGAGAGGTGGGGGCTGTTGAGACAGACG 1200
DB 1141 TGGTGTGTGTACTGCTCCAGCAGCAGCTGAGAGGTGGGGGCTGTTGAGACAGACG 1200
QY 1201 TCCAGGACCACTGAGTGTATGCTTACCTGCTGAGTCTGATGCAACCAAGCCG 1260
DB 1201 TCCAGGACCACTGAGTGTATGCTTACCTGCTGAGTCTGATGCAACCAAGCCG 1260
QY 1261 GGGGCGCGCGCTCCAGCCCGCCAAACAGAGAGGTTGTGCTCTTGGCGGAGAT 1320
DB 1261 GGGGCGCGCGCTCCAGCCCGCCAAACAGAGAGGTTGTGCTCTTGGCGGAGAT 1320

QY 1321 GGGCTCTGGAATCAGAAAATCCTATTTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGAC 1380
DB 1321 GGGCTCTGGAATCAGAAAATCCTATTTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGAC 1380
QY 1381 GGGGAAGACGCTCTGCTTCTCAACATGAACATCTTCAGAAAGGACATCAACTGTGAG 1440
DB 1381 GGGGAAGACGCTCTGCTTCTCAACATGAACATCTTCAGAAAGGACATCAACTGTGAG 1440
QY 1441 AGGTACTACAGCTTCACTTCTGAGTTTTCAGAAATTTCTTTGAGCTATGTACTATATC 1500
DB 1441 AGGTACTACAGCTTCACTTCTGAGTTTTCAGAAATTTCTTTGAGCTATGTACTATATC 1500
QY 1501 CTGACGAGGGGAGGCGGGCCAGCCAGACAGGACGTGACAGGCTGTTGACCGAG 1560
DB 1501 CTGACGAGGGGAGGCGGGCCAGCCAGACAGGACGTGACAGGCTGTTGACCGAG 1560
QY 1561 TACCGGTTTCTGAAAGGAGCTTCTGGCACTCACCAGCCGCTTCTTGTGAGTCTCTG 1620
DB 1561 TACCGGTTTCTGAAAGGAGCTTCTGGCACTCACCAGCCGCTTCTTGTGAGTCTCTG 1620
QY 1621 AACGAGGAGACGAGGACCACTCGAAGAGTCTCTGCTGGAAGGTCTCGCCGACATC 1680
DB 1621 AACGAGGAGACGAGGACCACTCGAAGAGTCTCTGCTGGAAGGTCTCGCCGACATC 1680
QY 1681 AAGATGACCTGTTGCAAGTGGATCCAAAGCAAAGCTCAGAGCGACGGCTCCACCTGCG 1740
DB 1681 AAGATGACCTGTTGCAAGTGGATCCAAAGCAAAGCTCAGAGCGACGGCTCCACCTGCG 1740
QY 1741 CAGGCTCTTGGAGTTCTTTCAGCTGTTGTACAGATCCAGGAGGAGGATTTATCCAG 1800
DB 1741 CAGGCTCTTGGAGTTCTTTCAGCTGTTGTACAGATCCAGGAGGAGGATTTATCCAG 1800
QY 1801 CAGGCTCTGAGCCACTTCCAGGTGATCGTGTGAGCAATTTGCTTCAAAGTGGAGCAC 1860
DB 1801 CAGGCTCTGAGCCACTTCCAGGTGATCGTGTGAGCAATTTGCTTCAAAGTGGAGCAC 1860
QY 1861 ATGTCTCTGCTGTTGTTGAAAGCGCTGAGGAGCGCCAGGCTGCTGCACTTGTATGCG 1920
DB 1861 ATGTCTCTGCTGTTGTTGAAAGCGCTGAGGAGCGCCAGGCTGCTGCACTTGTATGCG 1920
QY 1921 GCCACCTACAGCGGAGGAGAGACCGCGAGGTGCTCCGAGGAGCGCACACGCTG 1980
DB 1921 GCCACCTACAGCGGAGGAGAGACCGCGAGGTGCTCCGAGGAGCGCACACGCTG 1980
QY 1981 TTGTGTGAGCTCAGACAGAGAGGACCTTCTGCTGAGAGCTACAGTGAACATCTGGCA 2040
DB 1981 TTGTGTGAGCTCAGACAGAGAGGACCTTCTGCTGAGAGCTACAGTGAACATCTGGCA 2040
QY 2041 GCGGCTCTGTGCAACAATCCAAACCTGATAGAGTGTCTGTACCCGAAATGCGCTGGGC 2100
DB 2041 GCGGCTCTGTGCAACAATCCAAACCTGATAGAGTGTCTGTACCCGAAATGCGCTGGGC 2100
QY 2101 AGCCGGGGGTGAGCTGCTGTCAAGGACTCAGACACCCCAACTGCAAACTTCAGAAC 2160
DB 2101 AGCCGGGGGTGAGCTGCTGTCAAGGACTCAGACACCCCAACTGCAAACTTCAGAAC 2160
QY 2161 CTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTGCGAGGACCTCTCTGCACTCTC 2220
DB 2161 CTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTGCGAGGACCTCTCTGCACTCTC 2220
QY 2221 ATAGCCAAATAGAATTTTGAAGGATGGAATCTCAGTGGCAACCGGTTGATTTCCAGGC 2280
DB 2221 ATAGCCAAATAGAATTTTGAAGGATGGAATCTCAGTGGCAACCGGTTGATTTCCAGGC 2280
QY 2281 ATGATGCTGCTTTCGAGGCGCTCGGCATCCCCAGTGCAGCTGATGATTCAGTTG 2340
DB 2281 ATGATGCTGCTTTCGAGGCGCTCGGCATCCCCAGTGCAGCTGATGATTCAGTTG 2340
QY 2341 AGGAAGTGTACAGCTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTCTCGGCAACCA 2400
DB 2341 AGGAAGTGTACAGCTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTCTCGGCAACCA 2400
QY 2401 CCACATCTGTTGATGAGGACCTCAGAGGAATGCACTGGAGGATTTGGGCTGAGGTTA 2460

QY	829	CACGATCTCAGGAGCCCTGTGTGCTCTGTGGAGGAGAAACGGCCACCGAGCTGCTT	888
Db	1130	CACGATCTCAGGAGCCCTGTGTGCTCTGTGGAGGAGAAACGGCCACCGAGCTGCTT	1189
QY	889	CTTAACAGCTTAATTCGGAGAGAGCTGCTCCTGAGCTATCTTTGCTCATCACACACGG	948
Db	1190	CTTAACAGCTTAATTCGGAGAGAGCTGCTCCTGAGCTATCTTTGCTCATCACACACGG	1249
QY	949	CCCAACGGCTTTGGAGAGCTCCACCGTCTGTGGAGACCCACGACATGTGGAGATCTG	1008
Db	1250	CCCAACGGCTTTGGAGAGCTCCACCGTCTGTGGAGACCCACGACATGTGGAGATCTG	1309
QY	1009	GGCTTCTCTGAGGAGAGAGAGAAATPACTTCTCAAGTATTTCCAAATGCGAGAGCAG	1068
Db	1310	GGCTTCTCTGAGGAGAGAGAGAAATPACTTCTCAAGTATTTCCAAATGCGAGAGCAG	1369
QY	1069	GCAGGCAAGTCTTCAATACGTGAGGAGACACAGAGCTCTCTTCAACATGTCTCGTC	1128
Db	1370	GCAGGCAAGTCTTCAATACGTGAGGAGACACAGAGCTCTCTTCAACATGTCTCGTC	1429
QY	1129	CCCTGTGTGCTGGGTGGTGTACCTGCTCCAGAGAGAGCTGGAGGGTGGGGGCTG	1188
Db	1430	CCCTGTGTGCTGGGTGGTGTACCTGCTCCAGAGAGAGCTGGAGGGTGGGGGCTG	1489
QY	1189	TTGAGACAGAGCTCCAGAGACCACTGCGAGTGTACATGCTCTACCTGCTGAGTCTGATG	1248
Db	1490	TTGAGACAGAGCTCCAGAGACCACTGCGAGTGTACATGCTCTACCTGCTGAGTCTGATG	1549
QY	1249	CAACCCAGCGGGGGCCCGCGCTCCAGCCCGCCACCCACAGAGAGGGTGTGCTCC	1308
Db	1550	CAACCCAGCGGGGGCCCGCGCTCCAGCCCGCCACCCACAGAGAGGGTGTGCTCC	1609
QY	1309	TTGGCGCAGATGGGCTCTGGAATCAGAAATCTATTTGAGGAGCAGGAGCTCCGGAAG	1368
Db	1610	TTGGCGCAGATGGGCTCTGGAATCAGAAATCTATTTGAGGAGCAGGAGCTCCGGAAG	1669
QY	1369	CAGGGCTAGACGGGGAAGAGCTCTGCTTCTCTCAACATGAACATCTTCCAGAGGAC	1428
Db	1670	CAGGGCTAGACGGGGAAGAGCTCTGCTTCTCTCAACATGAACATCTTCCAGAGGAC	1729
QY	1429	ATCAACTGTGAGAGGTACTACAGTCTCATCCACTTGAGTTTCCAGGAAATCTTTGACGCT	1488
Db	1730	ATCAACTGTGAGAGGTACTACAGTCTCATCCACTTGAGTTTCCAGGAAATCTTTGACGCT	1789
QY	1489	ATGTAATCTATCTCTGAGAGGGGAGGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGG	1548
Db	1790	ATGTAATCTATCTCTGAGAGGGGAGGGGGAGGGGAGGGGAGGGGAGGGGAGGGGAGGGG	1849
QY	1549	CTGTTGACCGAGTACGGTCTTTCTGAAAGAGCTTCTGGACATCACAGCGCTTCTG	1608
Db	1850	CTGTTGACCGAGTACGGTCTTTCTGAAAGAGCTTCTGGACATCACAGCGCTTCTG	1909
QY	1609	TTTGGACTCTCTGAACGAGGAGACACAGGAGCCACTCTGGAGAGAGTCTCTGCTGGAAGTC	1668
Db	1910	TTTGGACTCTCTGAACGAGGAGACACAGGAGCCACTCTGGAGAGAGTCTCTGCTGGAAGTC	1969
QY	1669	TCGCCGCACATCAAGATGGAACCTTTGAGTGGATCCAAAGCAAAGCTCAGAGCGAGGC	1728
Db	1970	TCGCCGCACATCAAGATGGAACCTTTGAGTGGATCCAAAGCAAAGCTCAGAGCGAGGC	2029
QY	1729	TCCACCTCTGAGCAGGCTCTTGGAGTCTTACGTGCTTGTACGAGATCCAGAGGAG	1788
Db	2030	TCCACCTCTGAGCAGGCTCTTGGAGTCTTACGTGCTTGTACGAGATCCAGAGGAG	2089
QY	1789	GAGTTTATCCAGAGGCTCTGAGCCTTCCAGGTGATCTGGTGCAGCAATTTGCTCC	1848
Db	2090	GAGTTTATCCAGAGGCTCTGAGCCTTCCAGGTGATCTGGTGCAGCAATTTGCTCC	2149
QY	1849	AAGATGAGCAGATGCTCTCTGTTCTGTTGAGCGCTGACAGGCGCCAGGCTGCTG	1908
Db	2150	AAGATGAGCAGATGCTCTCTGTTCTGTTGAGCGCTGACAGGCGCCAGGCTGCTG	2209

QY	1909	CACTTGTATGCGCCACCTACAGCGCGAGACCGGGAAGACCGCGAGGTGCTCCGAGGA	1968
Db	2210	CACTTGTATGCGCCACCTACAGCGCGAGACCGGGAAGACCGCGAGGTGCTCCGAGGA	2269
QY	1969	GCGCACACGCTTTGTTGGTGCAGCTCAGACACAGAGAGGACCGTTCTGCTGGAGCGCTACAGT	2028
Db	2270	GCGCACACGCTTTGTTGGTGCAGCTCAGACACAGAGAGGACCGTTCTGCTGGAGCGCTACAGT	2329
QY	2029	GAAACATCTGGCAGCGCGCTCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA	2088
Db	2330	GAAACATCTGGCAGCGCGCTCTGTGACCAATCCAAACCTGATAGAGCTGCTCTGTACCGA	2389
QY	2089	AATGCCCTGGGACCGCGGGGTGAACTGCTCTGTCAAGGACTCAGACACCCCAACTGC	2148
Db	2390	AATGCCCTGGGACCGCGGGGTGAACTGCTCTGTCAAGGACTCAGACACCCCAACTGC	2449
QY	2149	AAACTTCAGAACCTTGAGGCTGAAAGAGTGCAGATCTCCAGCTCAGCTGCGAGGACCTC	2208
Db	2450	AAACTTCAGAACCTTGAGGCTGAAAGAGTGCAGATCTCCAGCTCAGCTGCGAGGACCTC	2509
QY	2209	TCTCAGCTCTCATAGCCAATAAGAAATTTGACAAAGGATGGATCTCAGTGGCAACGGCGTT	2268
Db	2510	TCTCAGCTCTCATAGCCAATAAGAAATTTGACAAAGGATGGATCTCAGTGGCAACGGCGTT	2569
QY	2269	GGATTCACGAGATGATGCTGCTTTGGAGGGCTGCGGATCCCAAGTGCAGCTGCAG	2328
Db	2570	GGATTCACGAGATGATGCTGCTTTGGAGGGCTGCGGATCCCAAGTGCAGCTGCAG	2629
QY	2329	ATGATTCAGTTGAGGAGTGTGCTGAGTTCGGGCTTGTCTAGGAGATGGCTTCTGTG	2388
Db	2630	ATGATTCAGTTGAGGAGTGTGCTGAGTTCGGGCTTGTCTAGGAGATGGCTTCTGTG	2689
QY	2389	CTCGGCAACCAACCATCTGTTGAGTTCGACCTGACAGGAAATGCACTCGAGGATTTG	2448
Db	2690	CTCGGCAACCAACCATCTGTTGAGTTCGACCTGACAGGAAATGCACTCGAGGATTTG	2749
QY	2449	GGCTGAGGTTTATATGCGCAGGAGCTGAGGCAACCGTCTGACAGTACGAGCTTTGFG	2508
Db	2750	GGCTGAGGTTTATATGCGCAGGAGCTGAGGCAACCGTCTGACAGTACGAGCTTTGFG	2809
QY	2509	CTGAAGATCTGCGGCTCAGTCTGCTGCTGACAGCTGCGCTCAACTCTCAGTGTG	2568
Db	2810	CTGAAGATCTGCGGCTCAGTCTGCTGCTGACAGCTGCGCTCAACTCTCAGTGTG	2869
QY	2569	AACAGAGCTGAGAGAGCTGGAACCTGAGCTGAGTGGGGGAGCTCGGGTGTGCTG	2628
Db	2870	AACAGAGCTGAGAGAGCTGGAACCTGAGCTGAGTGGGGGAGCTCGGGTGTGCTG	2929
QY	2629	CTGCTGTGAGGCGCTCAGGCAATCCCAAGCTCCAGACCTCGCGTTGGGCAATC	2688
Db	2930	CTGCTGTGAGGCGCTCAGGCAATCCCAAGCTCCAGACCTCGCGTTGGGCAATC	2989
QY	2689	TGCGGCTGGGCTGCGGCTGAGGGTCTTCTGTTGCTGCTGAGGCTGCTGCTGCTG	2748
Db	2990	TGCGGCTGGGCTGCGGCTGAGGGTCTTCTGTTGCTGCTGAGGCTGCTGCTGCTG	3049
QY	2749	CTCGGAGCTGGAATTTGAGTTCAGACCTGCGAGCTGCGGCTGCTGCTGCTGCTG	2808
Db	3050	CTCGGAGCTGGAATTTGAGTTCAGACCTGCGAGCTGCGGCTGCTGCTGCTGCTG	3109
QY	2809	GAGGGCTGCAACATCCCGCTGAGACTCAGAACTGTGGTGTGATGCTGTGGGCTC	2868
Db	3110	GAGGGCTGCAACATCCCGCTGAGACTCAGAACTGTGGTGTGATGCTGTGGGCTC	3169
QY	2869	ACAGCAGGCTGTGAGAACTTTTACTTCCCTGGGATCAACAGACTTGTGACCGAC	2928
Db	3170	ACAGCAGGCTGTGAGAACTTTTACTTCCCTGGGATCAACAGACTTGTGACCGAC	3229
QY	2929	CTTTACTGACCAACACCGCTTAGGGGACACAGGTGTCCGACTGTTTTCAGACGGCTG	2988
Db	3230	CTTTACTGACCAACACCGCTTAGGGGACACAGGTGTCCGACTGTTTTCAGACGGCTG	3289
QY	2989	AGCCATCTGGCTGCAAACTCCGAGCTCTGTTTATTTGGGATGGACCTGAATAAATG	3048

Db 3290 AGCATCTGGCTGCAAACTCCGAGTCTCTGGTTATTGGATGGAGCTGAATAAATG 3349
 QY 3049 ACCACAGTAGTGTGGAGCGCTTCGAGTACAAAACCTATTATTTGGACATGGCTGCTGA 3108
 Db 3350 ACCACAGTAGTGTGGAGCGCTTCGAGTACAAAACCTATTATTTGGACATGGCTGCTGA 3409

RESULT 4
 BC028069
 LOCUS
 DEFINITION Homo sapiens NACHT, leucine rich repeat and PYD containing 12, mRNA
 (cDNA clone MGC:40117 IMAGE:5212737), complete cds.
 ACCESSION BC028069
 VERSION
 KEYWORDS
 SOURCE MGC.
 ORGANISM Homo sapiens (human)

REFERENCE
 AUTHORS
 1 (bases 1 to 3507)
 Strausberg, R.D., Feingold, B.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Sanchez, A., Whitting, M., Kerteman, M., Madan, A., Rodrigues, S.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
 MEDLINE
 PUBMED 22388257
 12477932
 2 (bases 1 to 3507)
 Strausberg, R.
 Direct Submission
 TITLE
 JOURNAL
 Submitted (08-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA

REMARK
 COMMENT
 NIH-MGC Project URL: <http://mgi.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc.mgc@nih.gov
 Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsurgeon, C., Vogt, J.B., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAK Plate: 62 Row: 9 Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21955153.

FEATURES
 source
 1. 3507
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:40117 IMAGE:5212737"
 /tissue_type="Blood, adult leukocytes"
 /clone_lib="NIH_MGC_118"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1. 3507
 /gene="NALP12"
 /note="synonyms: RNO2, PYPAF7"
 /db_xref="LocusID:91662"
 120..3305
 /codon_start=1
 /product="PYRIN-containing APAF1-like protein 7"
 /protein_id="AAH28069.1"
 /db_xref="GI:20380400"
 /db_xref="LocusID:91662"
 /translation="MLRTAGDGLCRSLTYLEELAEVLEKFKYVLGTATLGEKGIKIP
 WSGMERAGPLEMAQLLIHFGPEAWRLALSTFERINRKLWGERQEDVLDTPPG
 PSSLGNGSTCLLEVLTPRKDPQETVYVRKFRMLMEDNARLGEVNLSHRYTRL
 LLVKEHNPQVOQQLDITGRGHARTVGHQASPIKLTLPEDERPPPTVVMQGA
 AGICKMLAHKVMLDWADGKLFQGRFDYLFVINCENMQSATECMQDLIFCWPEKS
 APQLRLVPRLLIFIDGPELKPSPHDPQWPCWLCEWRKPELNLNLSLRKLLP
 ELSLLITTPALEKHLRHLRHEHVEILGFSEARKEIFYKIFHNAQAGQVFNVR
 RLEPPTMFCVLVWVCTCLOQLGEGGLLRQTSRTTAYMYLLSLMOPKPGAP
 NQPLPQNRGICSLAADLNQKILFEQDLRKHGLDGEVSALNNMIFQDKINCR
 YVSTHLSFQFAFYILDEGGAGPDQDVTLLTEYAFSERSFLATSRFLFGL
 LNESTRHLEKSLCWKYSPIHKDLOWIQSKASDSTLOQSLERFSCLYEIOEE
 FIQALSHPQIVVSNVSNASKMHWVSFCLKRCRQAQVHLHYGATYSDGEDRCSA
 GAHTLLVQLPERTVLDAYSEHLAALCTNPNLIELSLYNALRGSGVLLCCQGRHP
 NKLQNLRLKRKRISSEADLSAALIANRNLTRMDLNSGNGVGFMMLLCEGRHPV
 CLRLQNLKQLQESCAQEMASVLTGNPHLVLDLTGNALDELGLLCCGLRHPVC
 RLRLTLKICRLTAACDEASTLSVNOSRLDELISNELGDLVLLCEGLRHPCTCK
 LQTLRLGLCRGSAACGLSVVLOANHLRELDISENDLGDWGLWLLAEGLOHPACRL
 QKLWDDSGDLTKACACENDIFLTGINGLTIDYLLTNALGDTGVRLCKRLSHPGCKLR
 VLWLFGLMGLNTFTSRALAAKRVTPYLDIGC"
 153..395
 /note="PAAD DAPIN; Region: PAAD/DAPIN/Pyrin domain. This
 domain is predicted to contain 6 alpha helices and to have
 the same fold as the pfam00531 domain. This similarity may
 mean that this is a protein-protein interaction domain"
 /db_xref="CDD:pfam02758"
 753..1262
 /note="NACHT; Region: NACHT domain. This NTPase domain is
 found in apoptosis proteins as well as those involved in
 MHC transcription activation. This family is closely
 related to pfam00931"
 /db_xref="CDD:pfam05729"
 2268..3224
 /note="LRR_R1; Region: Leucine-rich repeats (LRRs),
 ribonuclease inhibitor (RI)-like subfamily. LRRs are 20-29
 residue sequence motifs present in many proteins that
 participate in protein-protein interactions and have
 different functions and cellular locations. LRRs
 correspond to structural units consisting of a beta strand
 (xxlxLxxN/CxL conserved pattern) and an alpha helix. This
 alignment contains 12 strands corresponding to 11 full
 repeats, consistent with the extent observed in the
 subfamily acting as Ran GTPase Activating Proteins
 (RanGAP1)"
 /db_xref="CDD:cd00116"

ORIGIN
 Query Match 82.9%; Score 2577; DB 9; Length 3507;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2817; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 289 GATCCCAGGAACCTACAGGACTATGTCCGAGGAAAATTCGGCTCATGGAAGACGC 348
Db 489 GATCCCAGGAACCTACAGGACTATGTCCGAGGAAAATTCGGCTCATGGAAGACGC 548
QY 349 AATCGCGCCTAGGGGAATGTCAACTCAGCCACCGGTACACCCGCTCCTCTGGTG 408
Db 549 AATCGCGCCTAGGGGAATGTCAACTCAGCCACCGGTACACCCGCTCCTCTGGTG 608
QY 409 AAGGAGCACTCAAAACCCCATGAGTCCAGCAGAGCTTCTGGACACAGGCCGGGACAC 468
Db 609 AAGGAGCACTCAAAACCCCATGAGTCCAGCAGAGCTTCTGGACACAGGCCGGGACAC 668
QY 469 GCGAGGACCGTGGACACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 528
Db 669 GCGAGGACCGTGGACACACAGGCTAGCCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 728
QY 529 GAGGAGCGCCCGAGCCGACCGCTGCTCATGCAAGCGCGGAGGATAGGCAAG 588
Db 729 GAGGAGCGCCCGAGCCGACCGCTGCTCATGCAAGCGCGGAGGATAGGCAAG 788
QY 589 TCCATGCTGGACACAAAGGTGATGCTGGACATGGCGGACGGGAAGCTTCCAAAGGACA 648
Db 789 TCCATGCTGGACACAAAGGTGATGCTGGACATGGCGGACGGGAAGCTTCCAAAGGACA 848
QY 649 TTTGATTATCTTCTACATCACTGAGGAGATGAACAGAGTGCACCGAATGACG 708
Db 849 TTTGATTATCTTCTACATCACTGAGGAGATGAACAGAGTGCACCGAATGACG 908
QY 709 ATCAAGACCTCATCTTACGCTGCTGCTGAGCCAGCGCTCTCCAGGAGCTCATC 768
Db 909 ATCAAGACCTCATCTTACGCTGCTGCTGAGCCAGCGCTCTCCAGGAGCTCATC 968
QY 769 CGAGTTCCCGAGCGCTCTTTCATCATCGAGCTTCGATGAGCTCAAGCCTTCTTTC 1028
Db 969 CGAGTTCCCGAGCGCTCTTTCATCATCGAGCTTCGATGAGCTCAAGCCTTCTTTC 1088
QY 829 CAGATCCTCAGGACCGCTGCTGCTGAGGAGGAGAAACGGCCCAAGAGCTGCTT 888
Db 1029 CAGATCCTCAGGACCGCTGCTGCTGAGGAGGAGAAACGGCCCAAGAGCTGCTT 1088
QY 889 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 948
Db 1089 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 1148
QY 949 CCACCGCTTTGAGAGCTCCACCGTCTGCTGAGCAACCCAGGCAATGGAGATCCTG 1008
Db 1149 CCACCGCTTTGAGAGCTCCACCGTCTGCTGAGCAACCCAGGCAATGGAGATCCTG 1208
QY 1009 GGCCTTCTGAGGACGAGGAGGATATCTTCAAGATATTTCCAAATGAGAGCAG 1068
Db 1209 GGCCTTCTGAGGACGAGGAGGATATCTTCAAGATATTTCCAAATGAGAGCAG 1268
QY 1069 GCGGGCAAGCTTCAATTACGTGAGGACCAAGAGCTCTTTCACCATGTGCTTCGTC 1128
Db 1269 GCGGGCAAGCTTCAATTACGTGAGGACCAAGAGCTCTTTCACCATGTGCTTCGTC 1328
QY 1129 CCGCTGGTGTGGTGTGTGTAACCTGCTCAGCAGAGCTGAGGAGTGGGGGCTG 1188
Db 1329 CCGCTGGTGTGGTGTGTGTAACCTGCTCAGCAGAGCTGAGGAGTGGGGGCTG 1388
QY 1189 TTGAGACAGAGCTCCAGGACCACTGAGTGTATGCTCTACCTGCTGAGTCTGATG 1248
Db 1389 TTGAGACAGAGCTCCAGGACCACTGAGTGTATGCTCTACCTGCTGAGTCTGATG 1448
QY 1249 CAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCAAGAGAGGTTGTGCTCC 1308
Db 1449 CAACCAAGCCGGGGCCCGCGCTCCAGCCCGCCCAACCAAGAGAGGTTGTGCTCC 1508
QY 1309 TTGGCGGACAGTGGGCTCTGGAATCAGAAAATCCTATTTGAGGAGCAGGACCTCCGGAAG 1368
Db 1509 TTGGCGGACAGTGGGCTCTGGAATCAGAAAATCCTATTTGAGGAGCAGGACCTCCGGAAG 1568
QY 1369 CAGCGCTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1428

Db 1569 CACGGCTAGACGGGGAAGAGCTCTCTGCTTCTCAACATGAACAATCTTCCAGAGGAC 1628
QY 1429 ATCAACTGTGAGGAGTACTACAGCTTCACTCACTTGTAGTTTCCAGGAAATCTTTTCAGCT 1488
Db 1629 ATCAACTGTGAGGAGTACTACAGCTTCACTCACTTGTAGTTTCCAGGAAATCTTTTCAGCT 1688
QY 1489 ATGTACTATATCTCTGGACGAGGGGAGCGGGGAGCCAGACCCAGAGAGCTGACACAGG 1548
Db 1689 ATGTACTATATCTCTGGACGAGGGGAGCGGGGAGCCAGACCCAGAGAGCTGACACAGG 1748
QY 1549 CTGTGTACCGAGTACCGCTTTTCTGAAAGAGCTTCTCTGGCACTCACCAGCCGCTTCTGTG 1608
Db 1749 CTGTGTACCGAGTACCGCTTTTCTGAAAGAGCTTCTCTGGCACTCACCAGCCGCTTCTGTG 1808
QY 1609 TTTGGACTCTCTGACGAGGAGACCCAGAGCCACCTGAGAGAGTCTCTGCTGGAAGGTC 1668
Db 1809 TTTGGACTCTCTGACGAGGAGACCCAGAGCCACCTGAGAGAGTCTCTGCTGGAAGGTC 1868
QY 1669 TCGCCGCACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGCGACGGC 1728
Db 1869 TCGCCGCACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAAGCTCAGAGCGACGGC 1928
QY 1729 TCCACCTCTGACAGAGGCTCCTTGGAGTCTTCACTGCTTGTACGAGATCCAGAGAGAG 1788
Db 1929 TCCACCTCTGACAGAGGCTCCTTGGAGTCTTCACTGCTTGTACGAGATCCAGAGAGAG 1988
QY 1789 GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGTGTGATCGTGTGACAGCAATTTGCCCTCC 1848
Db 1989 GAGTTTATCCAGAGGCGCTGAGCCACTTCCAGTGTGATCGTGTGACAGCAATTTGCCCTCC 2048
QY 1849 AAGTGGAGCACATGCTCTCTCTGTTCTGAAGCGCTGACAGAGCGCCAGGTCGTG 1908
Db 2049 AAGTGGAGCACATGCTCTCTCTGTTCTGTTCTGTAAGCGCTGACAGAGCGCCAGGTCGTG 2108
QY 1909 CACTTGTATGGCGCACCTTACAGCGGACCGGGAAGACCGCGGAGGTGCTCCGACGGA 1968
Db 2109 CACTTGTATGGCGCACCTTACAGCGGACCGGGAAGACCGCGGAGGTGCTCCGACGGA 2168
QY 1969 GCGCACACGCTGTTGTGTCAGCTCAGACAGAGAGGACCGTCTGCTGAGAGCTTACAGT 2028
Db 2169 GCGCACACGCTGTTGTGTCAGCT---ACCAGAGGAGCCGTTCTGTGAGAGCTACAGT 2225
QY 2029 GAACATCTGSCAGCGCCCTGTGCAACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA 2088
Db 2226 GAACATCTGSCAGCGCCCTGTGCAACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA 2285
QY 2089 AATGCCCTGGGCAAGCCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2148
Db 2286 AATGCCCTGGGCAAGCCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2345
QY 2149 AAATTTCAAGAACCTGAGGCTGAAGAGTGCCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2208
Db 2346 AAATTTCAAGAACCTGAGGCTGAAGAGTGCCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2405
QY 2209 TCTGAGCTCTCTATAGCCAATTAAGAAATTTGACAAGGATGATCTCAGTGGCAACCGGCTT 2268
Db 2406 TCTGAGCTCTCTATAGCCAATTAAGAAATTTGACAAGGATGATCTCAGTGGCAACCGGCTT 2465
QY 2269 GGAATTTCCAGGCAATGATGCTGCTTTTGGAGGGCTCGGGCATCCCAAGTGCAGGCTGCAG 2328
Db 2466 GGAATTTCCAGGCAATGATGCTGCTTTTGGAGGGCTCGGGCATCCCAAGTGCAGGCTGCAG 2525
QY 2329 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTG 2388
Db 2526 ATGATTCAGTTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTTCAGGAGATGGCTTCTGTG 2585
QY 2389 CTCGGCAACCAACCAATCTGTTGAGTTGGACCTGACAGGAAATGCACCTGGAGGATTTG 2448
Db 2586 CTCGGCAACCAACCAATCTGTTGAGTTGGACCTGACAGGAAATGCACCTGGAGGATTTG 2645
QY 2449 GGCCTGAGCTTACTATGCCAGGAGCTGAGGACCCAGTCTGACAGACTACCGACTTTGTGG 2508

Db	2646	GGCCTGAGGTTACTATGCCAGGAGCTGAGCCACCCAGTCTGCAGACTACGAGCTTTGTGG	2705
Qy	2509	CTGAAGATCTGCGGCTCACTGCTGCTGCTGAGAGCTGGCTCAACTCTCAGTGTG	2568
Db	2706	CTGAAGATCTGCGGCTCACTGCTGCTGCTGAGAGCTGGCTCAACTCTCAGTGTG	2765
Qy	2569	AACACAGACCTGAGAGAGCTGAGACCTGAGCCTGAATGAGCTGGGGACCTCGGGTGTG	2628
Db	2766	AACACAGACCTGAGAGAGCTGAGACCTGAGCCTGAATGAGCTGGGGACCTCGGGTGTG	2825
Qy	2629	CTGCTGTGAGGGCTCAGGATCCAGTCCAGCTCCAGACCTCGGGTGTGGGATC	2688
Db	2826	CTGCTGTGAGGGCTCAGGATCCAGTCCAGCTCCAGACCTCGGGTGTGGGATC	2885
Qy	2689	TGCGGGCTGGGCTGCGGCTGTCAGAGTCTTCTGCTGGTCTCCAGGCCAACACACAC	2748
Db	2886	TGCGGGCTGGGCTGCGGCTGTCAGAGTCTTCTGCTGGTCTCCAGGCCAACACACAC	2945
Qy	2749	CTCCGGAGCTGGACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGTTGCTGGCT	2808
Db	2946	CTCCGGAGCTGGACTTGAGTTTCAACGACCTGGGAGACTGGGGCTGTGTTGCTGGCT	3005
Qy	2809	GAGGGCTGCAACATCCGCTCGAGACTCCAGAACTGTGCTGGATGAGCTGTGGCTC	2868
Db	3006	GAGGGCTGCAACATCCGCTCGAGACTCCAGAACTGTGCTGGATGAGCTGTGGCTC	3065
Qy	2869	ACAGCAAGGCTGTGAGAACTCTTACTTTCACCTGGGATCAACACGACCTTGACCGAC	2928
Db	3066	ACAGCAAGGCTGTGAGAACTCTTACTTTCACCTGGGATCAACACGACCTTGACCGAC	3125
Qy	2929	CTTTACCTGACCAACACGCTTAGGGACACAGGTGTCGACTGCTTTGCAAGCGGTG	2988
Db	3126	CTTTACCTGACCAACACGCTTAGGGACACAGGTGTCGACTGCTTTGCAAGCGGTG	3185
Qy	2989	AGCCATCTGCTGCAAACTCCGAGTCTCTGTTATTTGGATGAGACCTGAATAAATG	3048
Db	3186	AGCCATCTGCTGCAAACTCCGAGTCTCTGTTATTTGGATGAGACCTGAATAAATG	3245
Qy	3049	ACCCACAGTGTGGAGCGCTTCGAGTAAACAAACCTTTATTTGGATGAGCTGCTGA	3108
Db	3246	ACCCACAGTGTGGAGCGCTTCGAGTAAACAAACCTTTATTTGGATGAGCTGCTGA	3305
RESULT 5			
LOCUS	AY116205	3563 bp mRNA linear PRI 21-MAY-2003	
DEFINITION	Homo sapiens monarch-1 splice form II mRNA, complete cds;		
ACCESSION	AY116205		
VERSION	AY116205.1 GI:21711822		
KEYWORDS	alternatively spliced.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1. (bases 1 to 3563)		
JOURNAL	Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.		
PUBMED	Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat		
REFERENCE	Protein that Controls Classical and Nonclassical MHC Class I Genes		
AUTHORS	J. Immunol. 170 (11), 5354-5358 (2003)		
TITLE	12759408		
JOURNAL	Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm		
PUBMED	Road, Chapel Hill, NC 27599, USA		
REFERENCE	Location/Qualifiers		
AUTHORS	1. .3563		
TITLE	/organism="Homo sapiens"		
JOURNAL	/mol_type="mRNA"		
FEATURES	/db_xref="taxon:9606"		
source	/chromosome="19"		
	/map="19q13.4"		

CDS	/cell_line="U937"		
	221..3241		
	/note="contains alternatively spliced exon 9"		
	/codon_start=1		
	/product="monarch-1 splice form II"		
	/protein_id="AA075143.1"		
	/db_xref="GI:21711822"		
	/translation="MLRTAGDGLCRSLTYLEEAVALKKPKLVLTATLGEKGKIP		
	WSEKAGPLMAQLLTHFGPEANRLALSTFERINRNDLWERQREDLVDRTPGG		
	PSSLNSTCLLEVSLLTPRDPQETRYDVRKFLMEDNRNARCECVNLSHRYRL		
	LLVKEHSPNQVQQLDTRGHARTVGHQASPIKIELTFEPEEPFPVTVVWQGA		
	AGIGKSLAHKMLMDADKLFQGRFDYLFYINCRMNQSAETCSQDLIFSCVEEP		
	APLOELIRVPERLLFIIDGDELKSPHDPQWCLWEKRPSTLQNLISLRKKLLP		
	ELSLITTPRTALEKLHLEHPRHEVLGFSEAEKVEFYVFNHQAQGOVNVVR		
	DNEPLTFMCFPLVCWVCTCLOQLEGGLLRQTSRTTAYMLYLLSLMQPKGAP		
	RLQPPNQRGLCSLAADGLMNQKILFEEQDLRKGDGEDVAFALNMNIFQKINCR		
	YYSITHSFQFFAAWYILDEGGAGPDQDVTLLRLLTEYAFSESFALALRSRFLFL		
	LNESTRHLEKSLKWPSPHIMDLLOWIQSKAQSDGTLQOQSLEFSCYIEIEE		
	FIQALSHFOVIVVNSIAKMEHVMVSSFLKRCRQAOLVLYLGYATYADGDRACSA		
	GAHTLLVLRPERTVLLDAYSEHLAAALCTNPNLIELSLYRNALSGRVKLLCOGLRH		
	PNCLONLRLKRCRISSEACEDLSNALIANLIRMDLSGNGVFPGMILLCEGLRHP		
	QCRQMIOQRKQLESAGACQEMASVLGTPHPLVELDLTGNALDELGLLCCGLRHPV		
	KLQTLRLKICRIGSAACRGLSVLQAHNLRELDSLNEIGLDGLVLLCGLRHPVC		
	LQKLWLFPGMDLNKMTSHRLAALRVTKPYLDTC"		
	869..892		
	/note="Region: P-loop"		
	2069..3238		
	/note="Region: leucine rich repeat"		
	221..289		
	/note="Region: pyrin domain"		
misc_feature			
misc_feature			
misc_feature			
ORIGIN			
Query Match	82.4%;	Score 2562;	DB 9; Length 3563;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2562;	Conservative	0; Mismatches	0; Indels 0; Gaps 0;
Qy	289	GATCCCCAGGAAACCTACAGGAGCTATCTCCGACGAAATTCGGCTCATGGAAGACCGC	348
Db	590	GATCCCCAGGAAACCTACAGGAGCTATCTCCGACGAAATTCGGCTCATGGAAGACCGC	649
Qy	349	AATGCGCCCTAGGGAAATGTCTCAACCTCAGCCACCGGTACACCGGCTCTCTGCTGGTG	408
Db	650	AATGCGCCCTAGGGAAATGTCTCAACCTCAGCCACCGGTACACCGGCTCTCTGCTGGTG	709
Qy	409	AAGGAGCACTAAACCCCATGAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC	468
Db	710	AAGGAGCACTAAACCCCATGAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC	769
Qy	469	GCGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGAC	528
Db	770	GCGAGGACCGTGGGACACAGGCTAGCCCTCAAGATAGAGACCTCTTTGAGCCAGAC	829
Qy	529	GAGGAGCCCGCGAGCCACCGCGACCGTGTGTCATGCAAGCGCGCGGAGGATAGGCAAG	588
Db	830	GAGGAGCCCGCGAGCCACCGCGACCGTGTGTCATGCAAGCGCGGAGGATAGGCAAG	889
Qy	589	TCCATGTGGGACACAAAGGTGATGCTGGCGCGACGGGAACTCTTCCAAAGGACAG	648
Db	890	TCCATGTGGGACACAAAGGTGATGCTGGCGCGACGGGAACTCTTCCAAAGGACAG	949
Qy	649	TTTGATTTATCTTCTTACATCACTGAGGAGATGAACACGAGTGCCACGGAATGCAGC	708
Db	950	TTTGATTTATCTTCTTACATCACTGAGGAGATGAACACGAGTGCCACGGAATGCAGC	1009
Qy	709	ATGCAAGACCTCATCTTTCAGCTGTGGCTCAGCCACGCGGCTCTCCAGGAGCTCATC	768
Db	1010	ATGCAAGACCTCATCTTTCAGCTGTGGCTCAGCCACGCGGCTCTCCAGGAGCTCATC	1069
Qy	769	CGAGTTCCCGAGCGCTCTTTCATCATCATCAAGCGCTTCGATGAGCTCAAGCTTCTTTC	828
Db	1070	CGAGTTCCCGAGCGCTCTTTCATCATCATCAAGCGCTTCGATGAGCTCAAGCTTCTTTC	1129

QY 829 CACGATCCTCAGGAGCCCTGGTGGCTTCTGCTGGAGAGAGAAACGGCCCAACGAGAGCTGCTT 888
Db 1130 CACGATCCTCAGGAGCCCTGGTGGCTTCTGCTGGAGAGAGAAACGGCCCAACGAGAGCTGCTT 1189
QY 889 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 948
Db 1190 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 1249
QY 949 CCCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCCGAGCATGTGGAGATCCTG 1008
Db 1250 CCCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCACCCCGAGCATGTGGAGATCCTG 1309
QY 1009 GGCTTCTCAGGAGAGAGAGAGAAATATCTTCAAGATATTTCCAAATGCGAGCAG 1068
Db 1310 GGCTTCTCAGGAGAGAGAGAGAAATATCTTCAAGATATTTCCAAATGCGAGCAG 1369
QY 1069 GCGGGCCCAAGTCTTCAATACGTGAGGAGCAACAGAGCTCTCTTCAACATGCTTCGTC 1128
Db 1370 GCGGGCCCAAGTCTTCAATACGTGAGGAGCAACAGAGCTCTCTTCAACATGCTTCGTC 1429
QY 1129 CCCTGGTGTGCTGGTGGTGTGTACTGCTGCTCCAGCAGAGCTGGAGGGTGGGGGCTG 1188
Db 1430 CCCTGGTGTGCTGGTGGTGTGTACTGCTGCTCCAGCAGAGCTGGAGGGTGGGGGCTG 1489
QY 1189 TTGAGACAGAGCTCCAGGAGCACCACTGCGAGTGTACATGCTTACCTGCTGAGTCTGATG 1248
Db 1490 TTGAGACAGAGCTCCAGGAGCACCACTGCGAGTGTACATGCTTACCTGCTGAGTCTGATG 1549
QY 1249 CAACCCAAAGCCGGGGCCCGCGCTCCAGCCGCCCAACAGAGAGAGGTTGTGCTCC 1308
Db 1550 CAACCCAAAGCCGGGGCCCGCGCTCCAGCCGCCCAACAGAGAGGTTGTGCTCC 1609
QY 1309 TTGGCGCAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 1368
Db 1610 TTGGCGCAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGCAGGACCTCCGGAAG 1669
QY 1369 CAGGCTTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1428
Db 1670 CAGGCTTAGACGGGGAAGAGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1729
QY 1429 ATCAACTGTGAGAGTACTACAGCTTCACTTCACTTGAATTTCCAGGAATCTTTGACAGCT 1488
Db 1730 ATCAACTGTGAGAGTACTACAGCTTCACTTCACTTGAATTTCCAGGAATCTTTGACAGCT 1789
QY 1489 ATGTACTATATCTTGAAGAGGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1548
Db 1790 ATGTACTATATCTTGAAGAGGGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1849
QY 1549 CTGTTGACAGAGTACGCTTTCTGAAAGAGAGCTTCTGGACATCACAGCGGCTTCTG 1608
Db 1850 CTGTTGACAGAGTACGCTTTCTGAAAGAGAGCTTCTGGACATCACAGCGGCTTCTG 1909
QY 1609 TTTGGACTCTGAAACGAGGAGACAGAGGACCACTGGAGAGAGTCTCTGCTGGAAGGTC 1668
Db 1910 TTTGGACTCTGAAACGAGGAGACAGAGGACCACTGGAGAGAGTCTCTGCTGGAAGGTC 1969
QY 1669 TCGCCGACATCAAGATGGAACCTTTCAGTGGATCCAAAGCAAAAGCTCAGAGCAGCGC 1728
Db 1970 TCGCCGACATCAAGATGGAACCTTTCAGTGGATCCAAAGCAAAAGCTCAGAGCAGCGC 2029
QY 1729 TCCACCTGACAGAGGCTCTTGGAGTCTTCACTGCTGTTGACAGATCCAGGAGGAG 1788
Db 2030 TCCACCTGACAGAGGCTCTTGGAGTCTTCACTGCTGTTGACAGATCCAGGAGGAG 2089
QY 1789 GAGTTATCCAGAGGCTCAGGACCTTCCAGTGTATCTGGTGCAGCAATTCGCTCC 1848
Db 2090 GAGTTATCCAGAGGCTCAGGACCTTCCAGTGTATCTGGTGCAGCAATTCGCTCC 2149
QY 1849 AAGATGAGCAGATGGTCTCTGTTCTGTCTGAAGCGCTGACAGAGGCGCCAGGTGCTG 1908
Db 2150 AAGATGAGCAGATGGTCTCTGTTCTGTCTGAAGCGCTGACAGAGGCGCCAGGTGCTG 2209
QY 1909 CACTTGTATGGGCCACCTACAGCGGAGCGGGAAGACCGCGGAGGTCTCCGAGGA 1968

Db 2210 CACTTGTATGGGCCACCTACAGCGGGAACGGGGAAGACCGCGAGGTCTCCGAGGA 2269
QY 1969 GCGCACACAGCTGTTGGTGGAGCTCAGACAGAGAGAGACCGCTTCTGCTGGACGCTACAGT 2028
Db 2270 GCGCACACAGCTGTTGGTGGAGCTCAGACAGAGAGAGACCGCTTCTGCTGGACGCTACAGT 2329
QY 2029 GAACTCTGCGACGCGGCTTGTGCAACCAATCCAAACCTGATAGAGCTGTCTGTGACGA 2088
Db 2330 GAACTCTGCGACGCGGCTTGTGCAACCAATCCAAACCTGATAGAGCTGTCTGTGACGA 2389
QY 2089 AATGCCCTGGGAGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2148
Db 2390 AATGCCCTGGGAGCGGGGGTGAAGCTGCTCTGTCAAGGACTCAGACACCCCACTGC 2449
QY 2149 AACTTTCCAGAACCTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTCGAGGAGCCTC 2208
Db 2450 AACTTTCCAGAACCTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTCGAGGAGCCTC 2509
QY 2209 TCTGAGCTCTCATAGCAATAAAGATTTGCAAGGATGATCTCAGTGGCAACGGGCTT 2268
Db 2510 TCTGAGCTCTCATAGCAATAAAGATTTGCAAGGATGATCTCAGTGGCAACGGGCTT 2569
QY 2269 GGAATCCCGAGCATGATGCTTGTGAGGCGCTGCGCATCCCACTGAGGCTGAG 2328
Db 2570 GGAATCCCGAGCATGATGCTTGTGAGGCGCTGCGCATCCCACTGAGGCTGAG 2629
QY 2329 ATGATTCACTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTG 2388
Db 2630 ATGATTCACTGAGGAAGTGTGAGTGGAGTCCGGGCTTGTGAGGAGATGGCTTCTGTG 2689
QY 2389 CTGCGCAACCAACCAATCTGGTGTGAGTGGACCTGACAGAAATGCACTGGAGATTTG 2448
Db 2690 CTGCGCAACCAACCAATCTGGTGTGAGTGGACCTGACAGAAATGCACTGGAGATTTG 2749
QY 2449 GGCCTGAGGTTACTATGCCAGGAGCTGAGGACCCAGCTGCGAGCTACGAGCTTTG 2508
Db 2750 GGCCTGAGGTTACTATGCCAGGAGCTGAGGACCCAGCTGCGAGCTACGAGCTTTG 2809
QY 2509 CTGAAGATCTGCGGCTCACTGCTGCTGTGACAGCTGGCTCAACTCTCAGTGTG 2568
Db 2810 CTGAAGATCTGCGGCTCACTGCTGCTGTGACAGCTGGCTCAACTCTCAGTGTG 2869
QY 2569 AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGG 2628
Db 2870 AACAGAGCTGAGAGAGCTGAGCTGAGCTGAGCTGAGGAGGAGGAGGAGGAGGAGG 2929
QY 2629 CTGCTGTGAGGCGCTCAGGACATCCACGTCGAAGCTCCAGACCTCGCGTTCGGGATC 2688
Db 2930 CTGCTGTGAGGCGCTCAGGACATCCACGTCGAAGCTCCAGACCTCGCGTTCGGGATC 2989
QY 2689 TGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTCTGTGCTGCTCCAGGCGCAACCAAC 2748
Db 2990 TGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTCTGTGCTGCTCCAGGCGCAACCAAC 3049
QY 2749 CTGCGGAGCTGAGCTTTGAGTTCAACGACTGGGAGCTGGGCGCTGTTGCTGGCT 2808
Db 3050 CTGCGGAGCTGAGCTTTGAGTTCAACGACTGGGAGCTGGGCGCTGTTGCTGGCT 3109
QY 2809 GAGGCGTCAACATCCCGCTGCAGATCCAGAACTGTGG 2850
Db 3110 GAGGCGTCAACATCCCGCTGCAGATCCAGAACTGTGG 3151

RESULT 6
AY095146
LOCUS AY095146
DEFINITION Homo sapiens PYRIN-containing APAF1-like protein 7 mRNA, complete cds.
ACCESSION AY095146
VERSION AY095146.1 GI:21314906
KEYWORDS
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 3827)
Wang, L., Manji, G.A., Grenier, J.M., Al-Garawi, A., Merriam, S.,
Lora, J.M., Geddes, B.J., Briskin, M., DiStefano, P.S., and Bertin, J.
PYPAF7, a Novel PYRIN-containing Apaf1-like Protein That Regulates
Activation of NF-kappa B and Caspase-1-dependent Cytokine
Processing
J. Biol. Chem. 277 (33), 29874-29880 (2002)
JOURNAL 12019269
PUBMED 2 (bases 1 to 3827)
Bertin, J.
AUTHORS Direct Submission
TITLE Submitted (10-APR-2002) Inflammation, Millennium Pharmaceuticals
Inc., 45 Sidney Street, Cambridge, MA 02139, USA
JOURNAL
FEATURES
source
1..3827
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
177..3362
/note="PYPAF7; PYRIN-containing nucleotide-binding site/
leucine-rich repeat family member"
/codon_start=1
/product="PYRIN-containing APAF1-like protein 7"
/protein_id="AA18227.1"
/db_xref="GI:21314907"
translation="MLRTAGDDEGLRLSTYLEAEVLEKPKLYLGTATLGEKIP
WGSMEKAGLEMAQLLIIFGPEAWRLALSTFERINRKLWQRQDLVDRTPPG
PSSLGNQSTQVQVLLDTGSHARTVGHQASPKIEITLFPDEBPPEPTVMQGA
LLVKEHSPMVOQLDGTGSHARTVGHQASPKIEITLFPDEBPPEPTVMQGA
AGIKSMALHMLMDWAGKLFQGRFDYLTINRMNQSAFECMQDLIFSCWPEPS
APLOELIRTPALELPIDIGDELKPSFHDPOQWCLWEERKPTLELNLNLRKLLP
ELSLIITRPTLEKHLRHLRPHVEILGSEARKEYFYKFNHAGQOVFNVR
DNEPLTFMFCVWVCTCLOQEGGLRLRSTRTTAVYMLYLLSLMQPKGAP
RLOPPNORGHCSLAADGLWOKILFEEODLRKHLDGEDVSALNNMI FOKDLNCR
YSETHLSFQFFAAMYIILDEGGAGPDQVTRLLTEYAFSEFSFLIALSRFLGL
LNEETRSHLEKSVNSPHMLDLQWIKQAQSGSTLQOQSLFFSCLYEIQEE
FIQALSHFIVIVSNVSAHMHVSSFLKRCRQAQVHLHYGATYSADGEDRACSA
GAHTLVLPRTVLLDAYSLAALCNFNLI BLSLYRNALSGRGVLLKCOGLRHP
NCKLQNLRLKCRSSACEDLSAALANKNLTRMDLSNGVGFPMWLLCEGLRHPQ
CRLQWLOUKQOLGSAQCEMASVLGTNPLVELDITGNALDELGLRLCOGLRHPVC
RLRLTLWKICRLTAACDELASTLSVNSQLRELDLSLDELGLVLLCEGLRHPCK
LQTLRLGLCRGSAACEGLSVVLQANHLRELDLSFNDLGDWGLMLLEGLQHPACRL
QKLWDSGLLTAKACENLYFTLGINQTLTDLYLTNNALGDTGVRLLCRLSHPGCKLR
VLWLFMDLNKWTSHRLAALRVTKPYLDIG"

Query Match 79.6%; Score 2475; DB 9; Length 3827;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
1;
QY 289 GATCCCCAGGAAACCTACAGGACATATGTCCAGGAAATTCGGCTCATGGAAGACCGC 348
DB 546 GATCCCCAGGAAACCTACAGGACATATGTCCAGGAAATTCGGCTCATGGAAGACCGC 605
QY 349 AATGCGCGCTAGGGGAATGTGTAACCTCAGCCACCGGTACACCCGCTCTCTGCTG 408
DB 606 AATGCGCGCTAGGGGAATGTGTAACCTCAGCCACCGGTACACCCGCTCTCTGCTG 665
QY 409 AAGGAGCACTCAAAACCCATCAGCTCAGCAGCAGCTTCTGGACACAGGCGGGGACAC 468
DB 666 AAGGAGCACTCAAAACCCATCAGCTCAGCAGCAGCTTCTGGACACAGGCGGGGACAC 725
QY 469 GCGAGACCGTGGGACACAGGCTAGCCCATCAAGATAGACACCTCTTTGAGCCAGAC 528
DB 726 GCGAGACCGTGGGACACAGGCTAGCCCATCAAGATAGACACCTCTTTGAGCCAGAC 785
QY 529 GAGGAGCGCCCGAGCCACCGCGCACCGGTGTCATGCAAGGCGCGGCAAGGATAGCAAG 588
|||||

ORIGIN

DB 786 GAGGAGCGCCCGAGCCACCGCACCGCTGATGCTCATGCAAGCGCGCGCAGGATAGCAAG 845
QY 589 TCCATGCTGGCACACAAAGGTGATGCTGACCTGGCGGCGGAAAGCTCTTCCAAAGCACA 648
DB 846 TCCATGCTGGCACACAAAGGTGATGCTGACCTGGCGGCGGAAAGCTCTTCCAAAGCACA 905
QY 649 TTTGATTATCTCTTACATCAATCAATGCAAGGAGATGAACCAAGAGTGCCACGGAATGCAGC 708
DB 906 TTTGATTATCTCTTACATCAATCAATGCAAGGAGATGAACCAAGAGTGCCACGGAATGCAGC 965
QY 709 ATGCAAGACCTCATCTTTCAGCTGCTGGCTGAGCCAGCGGCCCTCTCCAGGAGTCAATC 768
DB 966 ATGCAAGACCTCATCTTTCAGCTGCTGGCTGAGCCAGCGGCCCTCTCCAGGAGTCAATC 1025
QY 769 CGAGTTCCCGAGCGCTCTTTTCATCATCAAGCGCTTCGATGAGCTCAAGCTCTCTTTC 828
DB 1026 CGAGTTCCCGAGCGCTCTTTTCATCATCAAGCGCTTCGATGAGCTCAAGCTCTCTTTC 1085
QY 829 CACGATCTCAGGAGACCTGCTGCTGCTGGGAGGAGAAACGGCCACCGAGCTGCTT 888
DB 1086 CACGATCTCAGGAGACCTGCTGCTGCTGGGAGGAGAAACGGCCACCGAGCTGCTT 1145
QY 889 CTTAAGACCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATCAACCAACG 948
DB 1146 CTTAAGACCTTAATTCGGAAGAGCTGCTCCTGAGCTATCTTTGCTCATCAACCAACG 1205
QY 949 CCCAGCGCTTTGGAGAGCTCCACCGTCTGCTGGGAGGAGCCCGAGGATGTGGAGATCTG 1008
DB 1206 CCCAGCGCTTTGGAGAGCTCCACCGTCTGCTGGGAGGAGCCCGAGGATGTGGAGATCTG 1265
QY 1009 GGCCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1068
DB 1266 GGCCTCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1325
QY 1069 GCGGCCCAAGTCTTCAATTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1128
DB 1326 GCGGCCCAAGTCTTCAATTAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1385
QY 1129 CCCCTGCTGCTGGGCTGGTGTGATGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1188
DB 1386 CCCCTGCTGCTGGGCTGGTGTGATGCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1445
QY 1189 TTGACAGACAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1248
DB 1446 TTGACAGACAGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1505
QY 1249 CAACCCAGCGCGGGGCGCGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1308
DB 1506 CAACCCAGCGCGGGGCGCGGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1565
QY 1309 TTGGCGGAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1368
DB 1566 TTGGCGGAGATGGGCTCTGGAATCAGAAATCCTATTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1625
QY 1369 CACGCTCTAGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1428
DB 1626 CACGCTCTAGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1685
QY 1429 ATCAACTGTGAGAGTACTACAGCTTTCATCACTTTGAGTTTCCAGGAATCTTTGAGCT 1488
DB 1686 ATCAACTGTGAGAGTACTACAGCTTTCATCACTTTGAGTTTCCAGGAATCTTTGAGCT 1745
QY 1489 ATGTAATATATCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1548
DB 1746 ATGTAATATATCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1805
QY 1549 CTGTTGACCGGAGTACCGCTTTTCTGAAAGAGCTTCTTGGCACTCACCAGCGCTTCTG 1608
DB 1806 CTGTTGACCGGAGTACCGCTTTTCTGAAAGAGCTTCTTGGCACTCACCAGCGCTTCTG 1865
QY 1609 TTTGAGCTCTTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
DB 1866 TTTGAGCTCTTGAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1925
|||||

FIQALSHFOVIVVNSIASKMHEVMSFSLKRCRSAOVLHLYGATYSADGEDRARCSEA
GAHTLLVOLPERTVLDAVSEHIAALATNPNLIELSLYRNALGSGRVKLLCOGLRHP
NKLONLRLEKCRISSSACEDLSAALLANKNLTEMDSLGNVGPGPMWMLCEGLRHPQ
CRLOMIQRKQLESAGCAQEWASVLGTNPPLHVELDLTGNLEDLGLRLLCOGLRHPVC
RLRTLWIKICRLTAAACDELAUSTLVNOSURELDLSLNEGLDGLVLLLCCEGLRHPCK
LQTLRLGICRLGSAAECGLSVLQANHLRELDLSFNDLGDWGLWLLAEGLQHPACRL
QKWLDSGTLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLLCKRLSHPGCKLR
VLWLFMDLKNMTHSRSLAALRVTKPYLDIGC"

ORIGIN

Query Match		79.6%; Score 2475; DB 9; Length 3827;
Best Local Similarity		99.8%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 2; Indels 3; Gaps 1;		
Qy	289	GATCCCAGGAAACCTACAGGACTATGTCCGCAAGGAATTCGGCTCATCGAAGACGC 348
Db	546	GATCCCAGGAAACCTACAGGACTATGTCCGCAAGGAATTCGGCTCATCGAAGACGC 605
Qy	349	AATCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCCGCTCTCTGTGGTG 408
Db	606	AATCGCGCTAGGGGAATGTCAACCTCAGCCACCGGTACACCCGCTCTCTGTGGTG 665
Qy	409	AAGGAGCACTCAAAACCCCATCGAGTTCAGCAGAGCTTCTTGACACAGGCCGGGAC 468
Db	666	AAGGAGCACTCAAAACCCCATCGAGTTCAGCAGAGCTTCTTGACACAGGCCGGGAC 725
Qy	469	GCGAGGACCGTGGGACACACAGCTAGCCATCAAGATAGAGACCTCTTTGAGCCAGAC 528
Db	726	GCGAGGACCGTGGGACACACAGCTAGCCATCAAGATAGAGACCTCTTTGAGCCAGAC 785
Qy	529	GAGGAGCGCCCGAGCCACCGCGCACCGTGTCTATGCAAGCGCGGCAAGGATAGGCAAG 588
Db	786	GAGGAGCGCCCGAGCCACCGCGCACCGTGTCTATGCAAGCGCGGCAAGGATAGGCAAG 845
Qy	589	TCCATGTGGCACACAAGGTGATGTGACCTGGGCGGACGGGAAGCTCTTCAAGGCGAGA 648
Db	846	TCCATGTGGCACACAAGGTGATGTGACCTGGGCGGACGGGAAGCTCTTCAAGGCGAGA 905
Qy	649	TTTGATTTCTTCTACATCACTGCGAGGAGATGAACAGAGTGGCCACCGAATGCAGC 708
Db	906	TTTGATTTCTTCTACATCACTGCGAGGAGATGAACAGAGTGGCCACCGAATGCAGC 965
Qy	709	ATGCAAGACCTCATCTTCAGCTGTGGCTCAGCCCGCGGCTCTCCAGAGCTCATC 768
Db	966	ATGCAAGACCTCATCTTCAGCTGTGGCTCAGCCCGGCTCTCCAGAGCTCATC 1025
Qy	769	CGAGTTCGCGAGCGCTCTTTTCAATCATCGACGGCTTCGATGAGCTCAAGCTTCCTTC 828
Db	1026	CGAGTTCGCGAGCGCTCTTTTCAATCATCGACGGCTTCGATGAGCTCAAGCTTCCTTC 1085
Qy	829	CACGATCTCAGGACCGCTGTGCTCTGCTGGGAGGAACCGCCACCGAGCTGCTT 888
Db	1086	CACGATCTCAGGACCGCTGTGCTCTGCTGGGAGGAACCGCCACCGAGCTGCTT 1145
Qy	889	CTTAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGG 948
Db	1146	CTTAACAGCTTAAATTCGGAAGAAGCTGCTCCCTGAGCTATCTTTGCTCATCACACAGG 1205
Qy	949	CCACGGCTTTGGAGAAGCTCCACCGTCTGTGGAGCACCCACAGCATGTGGAGATCTTG 1008
Db	1206	CCACGGCTTTGGAGAAGCTCCACCGTCTGTGGAGCACCCACAGCATGTGGAGATCTTG 1265
Qy	1009	GGCTTCTGAGGACAAAGGAGAAATCTTCTACAAGTATTTCCAAATGCAGAGCAG 1068
Db	1266	GGCTTCTGAGGACAAAGGAGAAATCTTCTACAAGTATTTCCAAATGCAGAGCAG 1325
Qy	1069	GCGGCGCAAGTCTCAATTTAGCTGAGGACAAACAGCTCTCTTCCACCATGTCTTCGTC 1128
Db	1326	GCGGCGCAAGTCTCAATTTAGCTGAGGACAAACAGCTCTCTTCCACCATGTCTTCGTC 1385
Qy	1129	CCCTGGTGTGTGGGTGTGTGTACCTGCCCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1188
Db	1386	CCCTGGTGTGTGGGTGTGTGTACCTGCCCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG 1445

Qy	1189	TTGAGACAGAGCTCCAGGACCAACACTGCAGTGTACATGCTTCTACCTGCTGAGTCTGATG 1248
Db	1446	TTGAGACAGAGCTCCAGGACCAACACTGCAGTGTACATGCTTCTACCTGCTGAGTCTGATG 1505
Qy	1249	CAACCCAAAGCGGGGGCCCGCGCTCTCAGCCCCCAACCCAAACAGAGAGGTTGTGTCC 1308
Db	1506	CAACCCAAAGCGGGGGCCCGCGCTCTCAGCCCCCAACCCAAACAGAGAGGTTGTGTCC 1565
Qy	1309	TTGGCGGAGATGGCTCTGGAATCAGAAATCTATTGAGGACAGGACTCCGGAG 1368
Db	1566	TTGGCGGAGATGGCTCTGGAATCAGAAATCTATTGAGGACAGGACTCCGGAG 1625
Qy	1369	CACGGCTTAGACGGGGAAGACGCTCTCGCTTCTCAACATGAACATCTTCCAGAAAGAC 1428
Db	1626	CACGGCTTAGACGGGGAAGACGCTCTCGCTTCTCAACATGAACATCTTCCAGAAAGAC 1685
Qy	1429	ATCAACTGTGAGAGTACTACAGCTTATCATCTGTAGTTCAGGAAATCTTTGACGCT 1488
Db	1686	ATCAACTGTGAGAGTACTACAGCTTATCATCTGTAGTTCAGGAAATCTTTGACGCT 1745
Qy	1489	ATGTACTATATCTTGGACGAGGGGAGGGGGGAGCCGAGACCCAGAGCTGACACAG 1548
Db	1746	ATGTACTATATCTTGGACGAGGGGAGGGGGGAGCCGAGACCCAGAGCTGACACAG 1805
Qy	1549	CTGTTGACCGAGTACGGCTTTCTGAAAGAGCTTCTTGGCACTCACCAGCGCTTCTCTG 1608
Db	1806	CTGTTGACCGAGTACGGCTTTCTGAAAGAGCTTCTTGGCACTCACCAGCGCTTCTCTG 1865
Qy	1609	TTTGAGCTCTTGAACAGGAGACACAGGAGCCACTGGAGAAAGTCTCTGTGGAAGTCT 1668
Db	1866	TTTGAGCTCTTGAACAGGAGACACAGGAGCCACTGGAGAAAGTCTCTGTGGAAGTCT 1925
Qy	1669	TCGCGGCACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGACGC 1728
Db	1926	TCGCGGCACATCAAGATGGACCTGTTGAGTGGATCCAAAGCAAGCTCAGAGCGACGC 1985
Qy	1729	TCCACCTGTGACGAGGGCTCTTGGAGTCTTTCAGCTCTTGTACGAGTCCAGAGGAG 1788
Db	1986	TCCACCTGTGACGAGGGCTCTTGGAGTCTTTCAGCTCTTGTACGAGTCCAGAGGAG 2045
Qy	1789	GAGTTTATCCAGCAGGGCTCAGCCACTTCAGAGTGTGCTGGTTCAGCAACATTCGCTCC 1848
Db	2046	GAGTTTATCCAGCAGGGCTCAGCCACTTCAGAGTGTGCTGGTTCAGCAACATTCGCTCC 2105
Qy	1849	AAGATGAGACATAGTCTCTCTGTTCTGTCTGAAGCTCAGAGCGCCAGGTGCTG 1908
Db	2106	AAGATGAGACATAGTCTCTCTGTTCTGTCTGAAGCTCAGAGCGCCAGGTGCTG 2165
Qy	1909	CACTTGTATGCGCCACCTACAGCGGACCGGGAAGACCGCGAGGTGCTCCGACAGA 1968
Db	2166	CACTTGTATGCGCCACCTACAGCGGACCGGGAAGACCGCGAGGTGCTCCGACAGA 2225
Qy	1969	GCGCACAGCTGTTGGTGCAGCTCAGACCAAGAGGACCGTTCTGCTGGACGCTTACAGT 2028
Db	2226	GCGCACAGCTGTTGGTGCAGCT--ACCAAGAGGACCGTTCTGCTGGACGCTTACAGT 2282
Qy	2029	GAAATCTGGCAGCGGCTGTGCACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2088
Db	2283	GAAATCTGGCAGCGGCTGTGCACCAATCCAACTGATAGAGCTGTCTCTGTACCGA 2342
Qy	2089	AATGCCCTGGGACGCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCAACTGC 2148
Db	2343	AATGCCCTGGGACGCGGGGGTGAAGCTGCTGTCAAGGACTCAGACACCCCAACTGC 2402
Qy	2149	AAATTTCAAGAACCTGAGGCTCAAGAGGTGCGGCATCTCCAGCTCAGCTCGGAGGACCTC 2208
Db	2403	AAATTTCAAGAACCTGAGGCTCAAGAGGTGCGGCATCTCCAGCTCAGCTCGGAGGACCTC 2462
Qy	2209	TCTGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGGATCTCAGTGGCAACCGGCTT 2268
Db	2463	TCTGAGCTCTCATAGCCAAATGAAGATTTGACAGGATGGATCTCAGTGGCAACCGGCTT 2522

QY	2269	GGATTCCAGGCGATGATGCTGCTTTTCGAGGGCCCTCGGGCATCCCAAGTCAGAGCTCGAG	2322
Db	2523	GGATTCCAGGCGATGATGCTGCTTTTCGAGGGCCCTCGGGCATCCCAATGCAAGCTCGAG	2582
QY	2329	ATGATTTCAGTTGAGGAAGTGTCACTCGAGTCCGGGGCTTGTGACGAGAGATGGCTTCGTG	2388
Db	2583	ATGATTTCAGTTGAGGAAGTGTCACTCGAGTCCGGGGCTTGTGACGAGAGATGGCTTCGTG	2642
QY	2389	CTGGGCAACCAACCCACATCTGGTTGAGTTGGACCTGCACAGGAATGCACCTGGAGGATTTG	2448
Db	2643	CTTGGCCACCAACCCACATCTGGTTGAGTTGGACCTGCACAGGAATGCACCTGGAGGATTTG	2702
QY	2449	GGCCTGAGGTTACTATGCCAGGAGCTGAGGGACCCAGTCTGCGAGCTTACCGAATTTGTGG	2508
Db	2703	GGCCTGAGGTTACTATGCCAGGAGCTGAGGGACCCAGTCTGCGAGCTTACCGAATTTGTGG	2762
QY	2509	CTGAAGATCTGCCGCTTCACTGCTGCTGCTGTGACGAGCTGGCCTCAAATCTCAGTGTG	2568
Db	2763	CTGAAGATCTGCCGCTTCACTGCTGCTGCTGTGACGAGCTGGCCTCAAATCTCAGTGTG	2822
QY	2569	AACAGAGCCTGAGAGAGCTGGACCTTGAGCCTGAAATGAGCTGGGGGACCTCGGGGTGCTG	2628
Db	2823	AACAGAGCCTGAGAGAGCTGGACCTTGAGCCTGAAATGAGCTGGGGGACCTCGGGGTGCTG	2882
QY	2629	CTGCTGTGTGAGGGCCTCAGGCATCCCAAGTCMAAGCTCCAGACCTTCGGTGTGGGCATC	2688
Db	2883	CTGCTGTGTGAGGGCCTCAGGCATCCCAAGTCMAAGCTCCAGACCTTCGGTGTGGGCATC	2942
QY	2689	TGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAAC	2748
Db	2943	TGCCGGCTGGGCTCTGCCGCTGTGAGGGTCTTTCTGTGTGCTCCAGGCCAACCAAC	3002
QY	2749	CTCGGGAGCTGGACTTTGAGTTTCAACGACCTGGGACACTGGGGCCTGTGGTTCGTGGCT	2808
Db	3003	CTCGGGAGCTGGACTTTGAGTTTCAACGACCTGGGACACTGGGGCCTGTGGTTCGTGGCT	3062
QY	2809	GAGGGGTGCAACATCCCGCTCGAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCCTC	2868
Db	3063	GAGGGGTGCAACATCCCGCTCGAGACTCCAGAAACTGTGGCTGGATAGCTGTGGCCTC	3122
QY	2869	ACAGCCAAGGCTGTGAGATCTTTACTTCACTTGGGATCAACAGACCTTTGACCGAC	2928
Db	3123	ACAGCCAAGGCTGTGAGATCTTTACTTCACTTGGGATCAACAGACCTTTGACCGAC	3182
QY	2929	CTTTACTGACCAACACGCGCTAGGGGACACAGGTGTCCGACTGTCTTTCGAAAGCGGCTG	2988
Db	3183	CTTTACTGACCAACACGCGCTAGGGGACACAGGTGTCCGACTGTCTTTCGAAAGCGGCTG	3242
QY	2989	AGCATCTCTGGCTGCAAACTCCGAGTCTCTGGTTATTTGGGATGGACCTGAAATAAATG	3048
Db	3243	AGCATCTCTGGCTGCAAACTCCGAGTCTCTGGTTATTTGGGATGGACCTGAAATAAATG	3302
QY	3049	ACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAACCTTATTTGGACATTCGCTGCTGA	3108
Db	3303	ACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAACCTTATTTGGACATTCGCTGCTGA	3362
RESULT	8		
LOCUS	AY116206	3395 bp mRNA linear PRI 21-MAY-2003	
DEFINITION	Homo sapiens monarch-1 splice form III mRNA, complete cds;		
ACCESSION	AY116206	alternatively spliced.	
VERSION	AY116206.1	GI:21711824	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
REFERENCE	1 (bases 1 to 3395)		
AUTHORS	Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.		
TITLE	Monarch-1: A Pyrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes		

J. Immunol.	170 (11),	5354-5358 (2003)
PUBLISHED	12/59408	
REFERENCE	2 (bases 1 to 3395)	
AUTHORS	Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.	
TITLE	Direct Submission	
JOURNAL	Submitted (29-May-2002) Linberger Cancer Center, UNC, Mason Farm Road, Chapel Hill, NC 27599, USA	
FEATURES	Location/Qualifiers	
source	1. .3395	
	/organism="Homo sapiens"	
	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
	/chromosome="19"	
	/map="19q13.4"	
	221. .3073	
CDS	/notes="contains alternatively spliced exons 7 and 8"	
	/codon_start=1	
	/product="monarch-1 splice form III"	
	/protein_id="AAW75144.1"	
	/db_xref="GI:21711825"	
	/translation="MLRTAGDGLCRLSTLYLEAEAVELKKFVLGTATLGEKGKIPWSGMEKAGLEMAQLLIITHFGEAWRLALSTFERINRKLWERCORDLVDRTPPGPSSLGNSQSTLLEVLSTPRKDPQRTYRDYVERKFLMEDNRARLGEVNLSHRYTRL LVLKSHSNPMQVQQLLDDTGRGHARTVGHQASPIKIELTFEPDERPPEPTVVMQGAAGIGKMLAHKMLWADGKLFQGRFDYLIYINCREMMSQATECSMDQLIFSCWPEPSAPLQELTIVPERLLFIIDGFLDLPXSFHDPPQWCLCWEERKPTLLNLNLSIRKLLPLLSLITTPRTALEKLRLHPHVEILGFSEARKEVYFYKHNAQAGQVFNVRDNEPLFTMCVPLVVCWVCTCLOQLEGGLRQTSRTTAVMYLLLSLMQPKDPA RLQPPNORGLCSLAADGLWNOKILFEBQDLRKHGLGDEYSAFLNMNIFQKDLNCE RYVSFHLSPQEPFAMYILIDBEGGAGDPQDVRTLLTEYAFSPERSFLATRSFLFGL LNEETSHLEKLSVKNWSPHIMDLLOLQSKAQSDVTLQOGLSEFFSCLYETQEE FTQALSHQFQVTVNSIASKMEHWSYFLCKRCQAQVHLHYGATYSADGBDRACSA GAHTLLVQLRPRTVLIDAYSLEHAALCTNPNLTLSLYRNALGSGVKLLICQGLRH FNCKLQNLKRCR.TSSACEDLSAALIANKNLITRMDLSGNGVPGPMWLLCEGLRHPCRLQWLQRLKQLESQAQENASVLGTPNPHLVELDLTGNALEDLGLRLICQGLRHYP CRLTLWLWLSGGLTAKACENLYFTLGINOTLTDLYLTNNALGDTGVRLLCRKLSHPGCKLRVLWLFPMGLMNRHSRLAURLVKPYLDIDGC"	
misc_feature	869. .892	
	/note="Region: P-loop"	
misc_feature	2069. .3070	
	/note="Region: leucine rich repeat"	
misc_feature	221. .289	
	/note="Region: pyrin domain"	
ORIGIN		
Query Match	71.5%;	Score 2223; DB 9; Length 3395;
Best Local Similarity	100.0%;	Pred. No. 0;
Matches 223;	Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	289	GATCCCAGGAACTTACAGGACATATGTCGCGAGAAATTCGGCTCATGGAAGACCGC 348
Db	590	GATCCCAGGAACTTACAGGACATATGTCGCGAGAAATTCGGCTCATGGAAGACCGC 649
QY	349	AATGCCGCTAGGGGAATGTCACACTCAGCCACCGGTACACCCGCTCTCTGCTGTG 408
Db	650	AATGCCGCTAGGGGAATGTCACACTCAGCCACCGGTACACCCGCTCTCTGCTGTG 709
QY	409	AGAGGACACTCAAAACCCCATGCGAGTCCAGCAGCAGCTTCTGGACACAGCGCGGACAC 468
Db	710	AGAGGACACTCAAAACCCCATGCGAGTCCAGCAGCAGCTTCTGGACACAGCGCGGACAC 769
QY	469	GGGAGACCGTGGGACACCGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 528
Db	770	GGGAGACCGTGGGACACCGGCTAGCCCATCAAGATAGAGACCTCTTTGAGCCAGAC 829
QY	529	GAGGAGCGCCCGAGCCACCGCGCACCGTGGTTCATGCAAGCGCGGACGATAGGCAAG 588
Db	830	GAGGAGCGCCCGAGCCACCGCGCACCGTGGTTCATGCAAGCGCGGACGATAGGCAAG 889
QY	589	TCCATGCTGGCACAAAGGTGATGCTGGCATGCGGCGGACGGGAAGCTCTTCCAGGCGAGA 648
Db	890	TCCATGCTGGCACAAAGGTGATGCTGGCATGCGGCGGACGGGAAGCTCTTCCAGGCGAGA 949

649 TTTGATTATCTTCTTACATCAACTGAGGAGATGAACAGAGTGCCACCGAATGCAGC 708
Db |||||
950 TTTGATTATCTTCTTACATCAACTGAGGAGATGAACAGAGTGCCACCGAATGCAGC 1009
Qy |||||
709 ATGAAGACCTCATCTTTCAGTCTGCTGAGCCAGCCGCGCTCTCCAGGAGCTCATC 768
Db |||||
1010 ATGAAGACCTCATCTTTCAGTCTGCTGAGCCAGCCGCGCTCTCCAGGAGCTCATC 1069
Qy |||||
769 CGAGTCCCGAGCGCTCTTTCATCATCAACGCGCTTCGATGAGCTCAAGCCTCTCTTC 828
Db |||||
1070 CGAGTCCCGAGCGCTCTTTCATCATCAACGCGCTTCGATGAGCTCAAGCCTCTCTTC 1129
Qy |||||
829 CACGATCCTCAGGACCCCTGCTGCTCTGCTGGAGAGAAACGCGCCACAGGAGCTGCTT 888
Db |||||
1130 CACGATCCTCAGGACCCCTGCTGCTCTGCTGGAGAGAAACGCGCCACAGGAGCTGCTT 1189
Qy |||||
889 CTTAAAGCTTAATTCGAAGAAGCTGCTCTGCTGAGCTATCTTTGCTCATCAACACGG 948
Db |||||
1190 CTTAAAGCTTAATTCGAAGAAGCTGCTCTGCTGAGCTATCTTTGCTCATCAACACGG 1249
Qy |||||
949 CCCAGCGCTTTGGAGAGCTCCACGCTGCTGCTGGAGACCCACGCGATGTGGAGATCCTG 1008
Db |||||
1250 CCCAGCGCTTTGGAGAGCTCCACGCTGCTGCTGGAGACCCACGCGATGTGGAGATCCTG 1309
Qy |||||
1310 GGCCTTCTCTGAGGAGAGAAAGAAATACCTTCTACAAAGTATTTCCAAATGCGAGAGCAG 1068
Db |||||
1069 GCGGCGCAAGTCTCAATTAAGTGAAGGACAAAGAGCTCTCTTCCACCATGTGCTTCGTC 1128
Qy |||||
1370 GCGGCGCAAGTCTCAATTAAGTGAAGGACAAAGAGCTCTCTTCCACCATGTGCTTCGTC 1429
Db |||||
1129 CCCCTGCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188
Qy |||||
1430 CCCCTGCTGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1489
Db |||||
1189 TTGAGACAGAGCTCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1248
Qy |||||
1490 TTGAGACAGAGCTCCAGGACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1549
Db |||||
1249 CAACCCAGCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1308
Qy |||||
1550 CAACCCAGCGGGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1609
Db |||||
1309 TTGGCGGAGATGGCTCTGGAATCAGAAATCCTATTGAGGAGAGGAGCTCCGGAAG 1368
Qy |||||
1610 TTGGCGGAGATGGCTCTGGAATCAGAAATCCTATTGAGGAGAGGAGCTCCGGAAG 1669
Db |||||
1369 CACGCGCTTAGACGGGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1428
Qy |||||
1670 CACGCGCTTAGACGGGAGAGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1729
Db |||||
1429 ATCAACTGTGAGAGGTAATACAGCTTTCATCACTGAGTTCCTGAGGAAATCTTTGAGCT 1488
Qy |||||
1730 ATCAACTGTGAGAGGTAATACAGCTTTCATCACTGAGTTCCTGAGGAAATCTTTTTCAGCT 1789
Db |||||
1489 ATGTACTATATCTTGAAGAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1548
Qy |||||
1790 ATGTACTATATCTTGAAGAGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1849
Db |||||
1549 CTGTTGACCGAGTAGCGCTTTTCTGAAGAGGAGCTTCTGCGGAGCTCAGGAGGAGGAGGAG 1608
Qy |||||
1850 CTGTTGACCGAGTAGCGCTTTTCTGAAGAGGAGCTTCTGCGGAGCTCAGGAGGAGGAGGAG 1909
Db |||||
1609 TTTGAGCTCCTGAACGAGGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1668
Qy |||||
1910 TTTGAGCTCCTGAACGAGGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1969
Db |||||
1669 TCGCGGAGATCAAGATGAGCTGCTGAGTGGATCCAAAGCAAGCTCAGAGGAGGAGGAGGAG 1728
Qy |||||
1970 TCGCGGAGATCAAGATGAGCTGCTGAGTGGATCCAAAGCAAGCTCAGAGGAGGAGGAGGAG 2029
Db |||||
1729 TCCACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1788
Qy |||||

2030 TCCACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2089
Db |||||
1789 GAGTTTATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1848
Qy |||||
2090 GAGTTTATCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2149
Db |||||
1849 AAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1908
Qy |||||
2150 AAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2209
Db |||||
1909 CACTTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1968
Qy |||||
2210 CACTTGTATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2269
Db |||||
1969 GCGCACAGGCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2028
Qy |||||
2270 GCGCACAGGCTGTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2329
Db |||||
2029 GAAATCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2088
Qy |||||
2330 GAAATCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2389
Db |||||
2089 AATGCGCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2148
Qy |||||
2390 AATGCGCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2449
Db |||||
2149 AATGCGCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2208
Qy |||||
2450 AATGCGCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2509
Db |||||
2209 TCTGAGCTCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2268
Qy |||||
2510 TCTGAGCTCTCATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2569
Db |||||
2269 GGAATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2328
Qy |||||
2570 GGAATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2629
Db |||||
2329 ATGATTCAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2388
Qy |||||
2630 ATGATTCAGTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2689
Db |||||
2389 CTGCGCACCAACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
Qy |||||
2690 CTGCGCACCAACCCACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2749
Db |||||
2449 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2508
Qy |||||
2750 GGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2809
Db |||||
2509 CTG 2511
Qy |||||
2810 CTG 2812
Db |||||

RESULT 9
AX116207
LOCUS
DEFINITION Homo sapiens monarch-1 splice form IV mRNA, complete cds;
ACCESSION AX116207
VERSION AX116207.1 GI:21711826
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 3221)
AUTHORS Williams, K.L., Taxman, D.J., Linhoff, M.W., Reed, W. and Ting, J.P.Y.
TITLE Monarch-1: A Pylrin/Nucleotide-Binding Domain/Leucine-Rich Repeat Protein That Controls Classical and Nonclassical MHC Class I Genes
JOURNAL J. Immunol. 170 (11), 5354-5358 (2003)

PUBMED 12759408
2 (bases 1 to 3221)
AUTHORS Williams, K.L., Linhoff, M.W. and Ting, J.P.Y.
TITLE Direct Submission
JOURNAL Submitted (29-MAY-2002) Lineberger Cancer Center, UNC, Mason Farm
Road, Chapel Hill, NC 27599, USA
FEATURES Location/Qualifiers
source 1..3221
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
221..2893
/note="contains alternatively spliced exons 7, 8 and 9"
/codon_start=1
/product="monarch-1 splice form IV"
/protein_id="AAW75145.1"
/db_xref="GI:21711827"
/translation="MLRTAGRDGLRLSTVLEBLEAVELKKFKLYLGTATELGEGKIP
WGSMEKAGPELEAQLLTHFGPEEARLALSTPERINRDLWERGQREDLVDPGG
PSSLGQSTCLLSVTPKDPQETRYDYRRKFRIMEDNRNARLGE CVNLSHYRL
LVKESNMQVQQLDITGRHARTVHQASPIKIEITLPEDEERPEPPRTVWMOA
AGTGMKLAHKVMDWADGKLFQGRFDYLY INCREMOSATSCMODLIFSCWPEPS
APQELIRVPERLFIIDGDELKPSFHPDQGWCLCWEKRPTELLNSLRKLLP
ELSLITTPALEKRLHLEHPRHVELLGFSEARKEYFYKIFHNAQGVQNFVR
DNEPLFTMCFVPLVWVCTCLOOQLEGGLLRQTSRTTAVTMYLALLSLMQPKGAP
RUPPNQGLCSLAADGLNKKILFEQDLRKHLDGEDVSAPLNINIFPKDINGER
YYSFHLFSQEFPAAMYIIDEDEGAGPQDVRLITTEYAFERSEFLALTSRFLGL
LNEETRSKLKSVKNSPHI KMDLQWIOSKQSDGSTLQOGLSFFSCLYIOEE
FIQALSHQVIVVSNASKMEHWSVFLKRCRQAQVHLHYGATYSAGEDRACSA
GAHTLVQVLPERTVLIDAYSEHLAALCTNPNIHLNLYRNALRGVKKLCCGLRH
PCKQLNLRKCRISSEACEDLSAALANKNTRMDLSNGVGFPGMMLLCEGLRHP
QCLQWIKKQKQCGAOCMASVLGTPNTHVELDLTGNALEDLGLRLCQGLRHPV
CRLTLWLFQMDLNWTHSRLEALRVTKPYLDIGC"
misc_feature 869..892
/note="Region: P-loop"
misc_feature 2089..2896
/note="Region: leucine rich repeat"
misc_feature 221..289
/note="Region: pyrin domain"
ORIGIN
Query Match 71.4%; Score 2220; DB 9; Length 3221;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 289 GATCCCCAGGAACTACAGGACTATGTCGAGGAAATTCGGCTCATGGAAGCCGC 348
DB 590 GATCCCGAGAACTACAGGACTATGTCGAGGAAATTCGGCTCATGGAAGCCGC 649
QY 349 AATGCGGCTAGGGAAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGCTGGT 408
DB 650 AATGCGGCTAGGGAAATGTCAACCTCAGCCACCGGTACACCGGCTCCTGCTGGT 709
QY 409 AAGGACACTCAACCCCACTGAGGTCCAGACAGCTTCTGACACAGCCGCGGACAC 468
DB 710 AAGGACACTCAACCCCACTGAGGTCCAGACAGCTTCTGACACAGCCGCGGACAC 769
QY 469 GCGAGACCGTGGGACACCAAGGCTAGCCCCATCAAGATAGACCTCTTTGAGCCAGAC 528
DB 770 GCGAGACCGTGGGACACCAAGGCTAGCCCCATCAAGATAGACCTCTTTGAGCCAGAC 829
QY 529 GAGGAGCGCCGAGCCACCGCCACCGTGGTCAAGAGCGCGGAGGATAGGCAAG 588
DB 830 GAGGAGCGCCGAGCCACCGCCACCGTGGTCAAGAGCGCGGAGGATAGGCAAG 889
QY 589 TCCATCTGSCACACAAGGTGATCTGGAATGCGGCGGAGGAGCTCTTCCAGGAGA 648
DB 890 TCCATCTGSCACACAAGGTGATCTGGAATGCGGCGGAGGAGCTCTTCCAGGAGA 949
QY 649 TTTGATTATCTTCTACATCAATCGAGGAGATGAACAGAGTGCACGGAATGCAGC 708

DB 950 TTTGATTATCTTCTACATCAACTGCAGGAGATGAACAGAGTGCACCGAATGCAGC 1009
QY 709 ATGCAGACCTCATCTTTAGCTGCTGCTGAGCCAGCGGCTCTTCCAGGAGCTCATC 768
DB 1010 ATGCAGACCTCATCTTTAGCTGCTGCTGAGCCAGCGGCTCTTCCAGGAGCTCATC 1069
QY 769 CGAGTTCCGAGCGCTCTTTTCATCATCGAGCGCTTTCAGATGAGCTCAAGCCCTTTTC 828
DB 1070 CGAGTTCCGAGCGCTCTTTTCATCATCGAGCGCTTTCAGATGAGCTCAAGCCCTTTTC 1129
QY 829 CAGATCTCAGGAGCGCTGCTGCTGCTGAGGAGAGAAACGGCCACGAGGAGCTCTT 888
DB 1130 CAGATCTCAGGAGCGCTGCTGCTGCTGAGGAGAGAAACGGCCACGAGGAGCTCTT 1189
QY 889 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 948
DB 1190 CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGG 1249
QY 949 CCACGCTTTGGAGAGCTTCCACCGTCTGCTGAGACACCCAGGCAATGTGGAGATCCTG 1008
DB 1250 CCACGCTTTGGAGAGCTTCCACCGTCTGCTGAGACACCCAGGCAATGTGGAGATCCTG 1309
QY 1009 GGCTTCTCAGGAGAGAGAGGAATATCTTCAAGTATTTCCAAATGACAGAGCAG 1068
DB 1310 GGCTTCTCAGGAGAGAGAGGAATATCTTCAAGTATTTCCAAATGACAGAGCAG 1369
QY 1069 GCGGGCCAACTCTTCAATACGTGAGGAGCAACAGAGCTCTTTCACCATGTGCTCGTC 1128
DB 1370 GCGGGCCAACTCTTCAATACGTGAGGAGCAACAGAGCTCTTTCACCATGTGCTCGTC 1429
QY 1129 CCCTGCTGCTGGTGGTGTGTACTGCTCAGGAGCAGCTGAGAGGTGGGGGCTG 1188
DB 1430 CCCTGCTGCTGGTGGTGTGTACTGCTCAGGAGCAGCTGAGAGGTGGGGGCTG 1489
QY 1189 TTGAGACAGACGCTCCAGGACCACTGAGTGTGATGCTCTACCTGCTGAGTCTCATG 1248
DB 1490 TTGAGACAGACGCTCCAGGACCACTGAGTGTGATGCTCTACCTGCTGAGTCTCATG 1549
QY 1249 CAACCCAAAGCGGGGCCCCCGGCTCCAGCCCCCAACCCAGAGAGGTTGTGCTCC 1308
DB 1550 CAACCCAAAGCGGGGCCCCCGGCTCCAGCCCCCAACCCAGAGAGGTTGTGCTCC 1609
QY 1309 TTGCGGCGAGATGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGACTCTCGGAAG 1368
DB 1610 TTGCGGCGAGATGGCTCTGGAATCAGAAATCTTATTTGAGGAGCAGGACTCTCGGAAG 1669
QY 1369 CACGCGCTAGACCGGGAAGACGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1428
DB 1670 CACGCGCTAGACCGGGAAGACGCTCTGCTCTCTCAACATGAACATCTTCCAGAGGAC 1729
QY 1429 ATCAACTGTGAGAGTACTACAGCTTTCATCACTTTCAGGAAATCTTTGCAAGCT 1488
DB 1730 ATCAACTGTGAGAGTACTACAGCTTTCATCACTTTCAGGAAATCTTTGCAAGCT 1789
QY 1489 ATGTAATATCTTGAACGAGGAGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1548
DB 1790 ATGTAATATCTTGAACGAGGAGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1849
QY 1549 CTGTTACCGAGTACCGGCTTTCTGAAAGGAGCTTCTGCGACTCACCAGCCGCTTCTG 1608
DB 1850 CTGTTACCGAGTACCGGCTTTCTGAAAGGAGCTTCTGCGACTCACCAGCCGCTTCTG 1909
QY 1609 TTTGGAATCTCTGAACGAGGAGCAGGAGCCACTGGAGAGAGTCTTGTCTGGAAGGTC 1668
DB 1910 TTTGGAATCTCTGAACGAGGAGCAGGAGCCACTGGAGAGAGTCTTGTCTGGAAGGTC 1969
QY 1669 TCGCCGCACATCAAGATGAGCTTGTGAGTGAATCCAAAGCAAGCTCAGAGCGGCGG 1728
DB 1970 TCGCCGCACATCAAGATGAGCTTGTGAGTGAATCCAAAGCAAGCTCAGAGCGGCGG 2029
QY 1729 TCCACCTCAGCAGGCGCTCTTTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG 1788
DB 2030 TCCACCTCAGCAGGCGCTCTTTGGAGTCTTTCAGCTGCTTGTACGAGATCCAGGAGGAG 2089

1474 CAACCCAGCGGGGCCCCGCTCCAGCCCCCACCACAGAGAGGGTGTGCTCC 1533
1309 TTGGGGGAGATGGCTCTGGAATCAGAAAATCTATTGAGGAGCAGGACTCCGGAAG 1368
1534 TTGGGGGAGATGGCTCTGGAATCAGAAAATCTATTGAGGAGCAGGACTCCGGAAG 1593
1369 CACGGCTTAGACGGGAGAGAGCTCTGCTCCCTTCTCAACATGAACATCTTCAGNAGGAC 1428
1594 CACGGCTTAGACGGGAGAGAGCTCTGCTCCCTTCTCAACATGAACATCTTCAGNAGGAC 1653
1429 ATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTGACGCT 1488
1654 ATCAACTGTGAGAGGTACTACAGCTTTCATCCACTTGTAGTTTCCAGGAATCTTTGACGCT 1713
1489 ATGTACTATATCTTGGACGAGGGAGAGGGCGGGGAGGCCCCAGACCAAGCAGTCAACGAG 1548
1714 ATGTACTATATCTTGGACGAGGGAGAGGGCGGGGAGGCCCCAGACCAAGCAGTCAACGAG 1773
1549 CTGTTGACCGAGTACGGGCTTTTCTGAAGGAGCTTCTGCGCACTCACCAGCGCTTCTCTG 1608
1774 CTGTTGACCGAGTACGGGCTTTTCTGAAGGAGCTTCTGCGCACTCACCAGCGCTTCTCTG 1833
1609 TTTGGAATCTCTGAACGAGGAGACCAAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC 1668
1834 TTTGGAATCTCTGAACGAGGAGACCAAGGAGCCACTGAGAGAGTCTCTGCTGGAAGGTC 1893
1669 TCGCGGCACATCAGATGGAGCTTGTGAGTGGATCCAAAGCAAGCTCAGAGCGAGCGC 1728
1894 TCGCGGCACATCAGATGGAGCTTGTGAGTGGATCCAAAGCAAGCTCAGAGCGAGCGC 1953
1729 TCACACCTGACGAGGAGCTCTGCGAGTCTTTCAGCTGCTGTGACGAGATCCAGGAGGAG 1788
1954 TCACACCTGACGAGGAGCTCTGCGAGTCTTTCAGCTGCTGTGACGAGATCCAGGAGGAG 2013
1789 GAGTTTATCCAGCAGGAGCCCTGAGCCACTTCCAGGTGATCTGCTGAGCAACATTCGCTCC 1848
2014 GAGTTTATCCAGCAGGAGCCCTGAGCCACTTCCAGGTGATCTGCTGAGCAACATTCGCTCC 2073
1849 AAGATGAGACATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
2074 AAGATGAGACATATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2133
1909 CACTTGTATGCGGCCACTACAGCGGAGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 1968
2134 CACTTGTATGCGGCCACTACAGCGGAGCGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 2193
1969 GCGCACACGCTGTGCTGAGCTCAGACCAAGAGGAGCCGCTTCTGCTGAGCGGCTACAGT 2028
2194 GCGCACACGCTGTGCTGAGCTCAGACCAAGAGGAGCCGCTTCTGCTGAGCGGCTACAGT 2250
2029 GAACATCTGCGAGCGGCGCTGTCGACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA 2088
2251 GAACATCTGCGAGCGGCGCTGTCGACCAATCCAAACCTGATAGAGTGTCTCTGTACCGA 2310
2089 AATGCTCTGCGAGCGGCGGAGTGAAGTCTCTGTCAGAGTCTCAGACCCCACTGC 2148
2311 AATGCTCTGCGAGCGGCGGAGTGAAGTCTCTGTCAGAGTCTCAGACCCCACTGC 2370
2149 AACTTTAGAACCTGAGGTGAAGAGTGTGCGCATCTCCAGCTCAGCTCAGGAGGAGCTC 2208
2371 AACTTTAGAACCTGAGGTGAAGAGTGTGCGCATCTCCAGCTCAGCTCAGGAGGAGCTC 2430
2209 TCTGAGCTCTCATAGCCCAATAGATTTGACAGAGATGATCTCAGTGGCAACGGGTT 2268
2431 TCTGAGCTCTCATAGCCCAATAGATTTGACAGAGATGATCTCAGTGGCAACGGGTT 2490
2269 GGAATCCAGGAGTATGCTGCTTTTCCAGGAGGCTCGGCGCATCCCGAGTGCAGGCTGCAG 2328
2491 GGAATCCAGGAGTATGCTGCTTTTCCAGGAGGCTCGGCGCATCCCGAGTGCAGGCTGCAG 2550
2329 ATGATTCAGTTCAGGAGTGTACGCTGAGTGTGCGGCTTGTGAGGAGATGGCTTCTGTG 2388
2551 ATGATTCAGTTCAGGAGTGTACGCTGAGTGTGCGGCTTGTGAGGAGATGGCTTCTGTG 2610

QY 2389 CTCGSCACCAACCCACATCTCTGTTGAGTTGACCTGACAGGAATGCACTGGAGATTG 2448
Db 2611 CTTGGSCCAACCCACATCTCTGTTGAGTTGACCTGACAGGAATGCACTGGAGATTG 2670
QY 2449 GGCCTGAGGTTTACTATGTCAGGAGCTGAGGACCCAGTCTCAGACTACGAGCTTTGTGG 2508
Db 2671 GGCCTGAGGTTTACTATGTCAGGAGCTGAGGACCCAGTCTCAGACTACGAGCTTTGTGG 2730
QY 2509 CTGAAGATCTCGCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
Db 2731 CTGAAGATCTCGCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2790
QY 2569 AACCAGAGCTTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2628
Db 2791 AACCAGAGCTTGAGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2850
QY 2629 CTGCTGTGTGAGGCGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTG 2679
Db 2851 CTGCTGTGTGAGGCGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTG 2901

RESULT 11
AK095460
LOCUS
DEFINITION
Homo sapiens cDNA FLJ38141 fis, clone D90ST2002673, weakly similar to Homo sapiens caspase recruitment domain protein 7 mRNA.
ACCESSION
AK095460
VERSION
AK095460.1 GI:21754717
KEYWORDS
oligo capping; fis (full insert sequence).
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
REFERENCE
1
AUTHORS
Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Ohshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3466)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .3466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="D90ST2002673"
/cell_type="CD34+ Cells"
/clone_lib="D90ST2"
/note="cloning vector: pME18SFL3-mRNA from CD34+ cells after 9-days ODF induction.-primary culture, CD34+ Cells"

FEATURES
source
1. .3466
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="D90ST2002673"
/cell_type="CD34+ Cells"
/clone_lib="D90ST2"
/note="cloning vector: pME18SFL3-mRNA from CD34+ cells after 9-days ODF induction.-primary culture, CD34+ Cells"
ORIGIN
Query Match 64.2%; Score 1995; DB 9; Length 3466;

	Best Local Similarity	99.7%; Pred. No. 0;	Mismatches	Conservative	0; Mismatches	3; Indels	3; Gaps	1;
QY	289	GATCCCCAGAAAACCTTACAGGACTATGTCCGCAGAAATTCCGGCTCATGAAGACCGC	348					
Db	514	GATCCCAGGAAAACCTA CAGGACTATGTCCGCAGGAANTTCGGCTCATGNAAGACCGC	573					
QY	349	AATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCAGGTACA CCGGGTCTCTGTGGTG	408					
Db	574	AATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCAGGTACA CCGGGTCTCTGTGGTG	633					
QY	409	AGGAGCACTCAAACCCCATGAGGTCAGCAGCAGCTTTCTGGACACAGGCCGGGAGCAC	468					
Db	634	AGGAGCACTCAAACCCCATGAGGTCAGCAGCAGCTTTCTGGACACAGGCCGGGAGCAC	693					
QY	469	GGGAGGACCGTGGGACAC CAGAGCTAGCCCCATCAAGATAGACACCTCTTTGAGCCAGAC	528					
Db	694	GGGAGGACCGTGGGACAC CAGAGCTAGCCCCATCAAGATAGACACCTCTTTGAGCCAGAC	753					
QY	529	GAGGAGCGCCCGAGCCAC CGGCACCGTGGTCA TCAAGGCGCGGACGGATAGGCAAG	588					
Db	754	GAGGAGCGCCCGAGCCAC CGGCACCGTGGTCA TCAAGGCGCGGACGGATAGGCAAG	813					
QY	589	TCCATGCTGGCACAAAGGTGATGCTGGACTGGCGGACGGGNAGCTCTTCCAAGGCAGA	648					
Db	814	TCCATGCTGGCACAAAGGTGATGCTGGACTGGCGGACGGGAGAGCTCTTCCAAGGCAGA	873					
QY	649	TTTGATTATCTCTTCTACATCAACTCAGGGAGATGAAC CCAGAGTGCCACGGAAATGCAGC	708					
Db	874	TTTGATTATCTCTTCTACATCAACTCAGGGAGATGAAC CCAGAGTGCCACGGAAATGCAGC	933					
QY	709	ATGCAAGACCTCATCTTTAGCTGCTGGCGCTGAGCCCGCAGCGCGCTCTCCAGGAGCTCATC	768					
Db	934	ATGCAAGACCTCATCTTTAGCTGCTGGCGCTGAGCCCGCAGCGCGCTCTCCAGGAGCTCATC	993					
QY	769	CGAGTTCGCGAGCGGCTCTTTTCATCATCGACGGCTTGGATGAGCTCAAGCCTTCTTTC	828					
Db	994	CGAGTTCGCGAGCGGCTCTTTTCATCATCGACGGCTTGGATGAGCTCAAGCCTTCTTTC	1053					
QY	829	CAGATCCTCAGGACCCCTGGTGCCTCTGCTGGAGGAGAAAA CGGCCACGGAGCTGCTT	888					
Db	1054	CAGATCCTCAGGACCCCTGGTGCCTCTGCTGGAGGAGAAAA CGGCCACGGAGCTGCTT	1113					
QY	889	CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGTCTCATCACCACAGG	948					
Db	1114	CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGTCTCATCACCACAGG	1173					
QY	949	CCCA CGGCTTTGAGAAAGCTCCACCGTCTGCTGGAGACACCCAGGCAATGGAAGATCCTG	1008					
Db	1174	CCCA CGGCTTTGAGAAAGCTCCACCGTCTGCTGGAGACACCCAGGCAATGGAAGATCCTG	1233					
QY	1009	GGCTTCTCTGAGGCAGAAAGGAGGATACTCTTACAAGTATTTCCACAATGACAGAGCAG	1068					
Db	1234	GGCTTCTCTGAGGCAGAAAGGAGGATACTCTTACAAGTATTTCCACAATGACAGAGCAG	1293					
QY	1069	GGGGCCAAAGCTTCAATTACGTGAGGGA CAACGAGCCTCTCTTCAACATGTGCTTCGTC	1128					
Db	1294	GGGGCCAAAGCTTCAATTACGTGAGGGA CAACGAGCCTCTCTTCAACATGTGCTTCGTC	1353					
QY	1129	CCCCTGGTGTCTGGGTGGTGTGPA CCTGCTCCAGCAGCAGCTGGAGGTTGGGGGCTG	1188					
Db	1354	CCCCTGGTGTCTGGGTGGTGTGPA CCTGCTCCAGCAGCAGCTGGAGGTTGGGGGCTG	1413					
QY	1189	TTGAGACAGAGCTCGAGGACCAACACATGACGTGTACATGCTCTACCTCTGAGTCTGATG	1248					
Db	1414	TTGAGACAGAGCTCGAGGACCAACACATGACGTGTACATGCTCTACCTCTGAGTCTGATG	1473					
QY	1249	CAACCCAGCCGGGGCCCCCGGCTCCAGCCCCCA ACCAACAGAGAGGGTTGTGCTCC	1308					
Db	1474	CAACCCAGCCGGGGCCCCCGGCTCCAGCCCCCA ACCAACAGAGAGGGTTGTGCTCC	1533					
QY	1309	TTGGCGGACAGATGGGCTCTGGAATCAGAAAAATCTATT TTGAGGACGAGGACCTCCGGAAG	1368					

[illegible]

[illegible]

QY 1189 TTGAGACAGAGCTCCAGGACACCACTGCACTGATCTCTACCTGCTGAGTCTGATG 1248
Db 79121 TTGAGACAGAGCTCCAGGACACCACTGCACTGATCTCTACCTGCTGAGTCTGATG 79062
QY 1249 CAACCCAAAGCGGGGCGCGGCTCCAGCCGCCAACCACAGAGAGGTTGTCTCC 1308
Db 79061 CAACCCAAAGCGGGGCGCGGCTCCAGCCGCCAACCACAGAGAGGTTGTCTCC 79002
QY 1309 TTGCGCGCAGATGGCTCTGGAATCAGAAATCTATTTCAGGAGCAGGACCTCCGGAAG 1368
Db 79001 TTGCGCGCAGATGGCTCTGGAATCAGAAATCTATTTCAGGAGCAGGACCTCCGGAAG 78942
QY 1369 CACGCGCTAGACGGGAGAGAGCTCTCTGCGCTTCCCAACATGAACATCTTCCAGAGGAC 1428
Db 78941 CACGCGCTAGACGGGAGAGAGCTCTCTGCGCTTCCCAACATGAACATCTTCCAGAGGAC 78882
QY 1429 ATCAACTGTGAGAGGTACTACAGCTTCATCCATCTGAGTTTCCAGGAATCTTTTGCAGCT 1488
Db 78881 ATCAACTGTGAGAGGTACTACAGCTTCATCCATCTGAGTTTCCAGGAATCTTTTGCAGCT 78822
QY 1489 ATGTACTATATCTCTGGAGCGGGAGGCGGCGGAGCCAGCCAGAGCTGACCAAG 1548
Db 78821 ATGTACTATATCTCTGGAGCGGGAGGCGGCGGAGCCAGCCAGAGCTGACCAAG 78762
QY 1549 CTGTTGACCGAGTACGGGTTTTCTGAAAGGAGCTTCTGCGCACTCACCAGCGCTTCTG 1608
Db 78761 CTGTTGACCGAGTACGGGTTTTCTGAAAGGAGCTTCTGCGCACTCACCAGCGCTTCTG 78702
QY 1609 TTTGGACTCTGAAACGAGGAGACAGGAGCCACTGGAAGAGTCTCTGCTGGAAGGTC 1668
Db 78701 TTTGGACTCTGAAACGAGGAGACAGGAGCCACTGGAAGAGTCTCTGCTGGAAGGTC 78642
QY 1669 TCGCCGACATCAAGATGGAGCTTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGGC 1728
Db 78641 TCGCCGACATCAAGATGGAGCTTTGTCAGTGGATCCAAAGCAAGCTCAGAGCGAGGC 78582
QY 1729 TCCACCTGACAGAGGCTCTCTGAGTCTTCTGAGTCTGTCAGTCTGTCAGATCCAGAGGAG 1788
Db 78581 TCCACCTGACAGAGGCTCTCTGAGTCTTCTGAGTCTGTCAGTCTGTCAGATCCAGAGGAG 78522
QY 1789 GAGTTATCCAGAGGCGCTGAGCCACTTCCAGGTGATCGTGGTACAGCAATTCGCTCC 1848
Db 78521 GAGTTATCCAGAGGCGCTGAGCCACTTCCAGGTGATCGTGGTACAGCAATTCGCTCC 78462
QY 1849 AAGATGGAGCATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1908
Db 78461 AAGATGGAGCATGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 78402
QY 1909 CACTTGTATGCGCCACCTACAGCGCGGAGCGGAGAGCCGCGGAGTGTCTCCGCGAGGA 1968
Db 78401 CACTTGTATGCGCCACCTACAGCGCGGAGCGGAGAGCCGCGGAGTGTCTCCGCGAGGA 78342
QY 1969 GCGCACACGCTGTTGGTGCAGCT 1991
Db 78341 GCGCACACGCTGTTGGTGCAGCT 78319

RESULT 13
AX575503
LOCUS Sequence 19 from Patent WO02072630.
DEFINITION
ACCESSION AX575503
VERSION AX575503.1 GI:27552095

KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Thornton, M., Hafalia, A.J., Lu, D.A., Arvizu, C., Swarnakar, A., Lu, Y.,
Warren, B.A., Baughn, M.R., Tang, Y.T., Lee, E.A., Yao, M.G.,
Ranumkar, J., Khan, F.A., Gandhi, A.R., Ding, L., Yue, H., Gietzen, K.J.,
Walia, N.K., Thangavelu, K., Elliot, V.S. and Marquis, J.P.

TITLE Nucleic acid-associated proteins
JOURNAL Patent: WO 02072630-A 19 SEP-2002;
Incyte Genomics, Inc. (US)
FEATURES
source location/Qualifiers
1. 2158
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 429930CB1"
ORIGIN
Query Match 43.6%; Score 1356; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 1408 ATGACATCTTCCAGAGGACATCACTGTGAGAGTACTACAGCTTCATCCACTTCAGT 1467
Db 1 ATGAACATCTTCCAGAGGACATCACTGTGAGAGTACTACAGCTTCATCCACTTCAGT 60
QY 1468 TTCCAGGAATCTTTCAGAGCTATGTAATATCTGACAGAGGAGGCGGCGGAGGC 1527
Db 61 TTCCAGGAATCTTTCAGAGCTATGTAATATCTGACAGAGGAGGCGGCGGAGGC 120
QY 1528 CCAGACAGAGAGCTGACAGGCTGTTGACGAGTACGCGTTTTCTGAAAGAGCTTCGT 1587
Db 121 CCAGACAGAGAGCTGACAGGCTGTTGACGAGTACGCGTTTTCTGAAAGAGCTTCGT 180
QY 1588 GCACTCACAGCGCTTCTTGGACTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAG 1647
Db 181 GCACTCACAGCGCTTCTTGGACTCTGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 1648 AAGAGTCTTCTGCGAAGGCTCTCGCGCACATCAAGATGAGACCTGTTGACGTGATCAA 1707
Db 241 AAGAGTCTTCTGCGAAGGCTCTCGCGCACATCAAGATGAGACCTGTTGACGTGATCAA 300
QY 1708 AGCAAGCTCAGAGCGAGCTCCACCTCGACAGAGGCTCTTGGAGTCTTTCAGGTGC 1767
Db 301 AGCAAGCTCAGAGCGAGCTCCACCTCGACAGAGGCTCTTGGAGTCTTTCAGGTGC 360
QY 1768 TTGTACAGATCCAGAGGAGGAGTATTCAGAGAGGCTTTCAGAGGCTTTCAGGTGATC 1827
Db 361 TTGTACAGATCCAGAGGAGGAGTATTCAGAGAGGCTTTCAGAGGCTTTCAGGTGATC 420
QY 1828 GTGTGTAGCAACATTTGCTCCAGATGAGACATGTTCTCTGCTGCTGCTGCTGCTGCTG 1887
Db 421 GTGTGTAGCAACATTTGCTCCAGATGAGACATGTTCTCTGCTGCTGCTGCTGCTGCTG 480
QY 1888 TGCAGAGCGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1947
Db 481 TGCAGAGCGCCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
QY 1948 CGCGGAGGTGCTCGCAGGAGCGCACAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2007
Db 541 CGCGGAGGTGCTCGCAGGAGCGCACAGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 597
QY 2008 GTTCTGTGAGCGCTTACAGTGAACATCTGCGAGCGGCGCTTGTGACCAATCCAAACCTG 2067
Db 598 GTTCTGTGAGCGCTTACAGTGAACATCTGCGAGCGGCGCTTGTGACCAATCCAAACCTG 657
QY 2068 ATAGAGCTGCTCTGTGACCGAAATGCCCTGGGAGCGGCGGCTGCTGCTGCTGCTGCTG 2127
Db 658 ATAGAGCTGCTCTGTGACCGAAATGCCCTGGGAGCGGCGGCTGCTGCTGCTGCTGCTG 717
QY 2128 GGACTCAGACACCCCACTGCAAACTTCAGAACTTCAGAGGCTGAGAGGTGCGGATCTCC 2187
Db 718 GGACTCAGACACCCCACTGCAAACTTCAGAACTTCAGAGGCTGAGAGGTGCGGATCTCC 777
QY 2188 AGCTCAGCTGCGAGGAGCTCTCTGAGCTCTCATAGCAATAAGATTTTCAGAGGATG 2247
Db 778 AGCTCAGCTGCGAGGAGCTCTCTGAGCTCTCATAGCAATAAGATTTTCAGAGGATG 837
2248 GATCTCAGTGCAACCGGCTTGGATTCAGAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTG 2307

Db	838	GATCTCAGTCGGCAACGGCGTTGGATTCCCAAGGCATGATGCTTTGCGAGGGCGCTCGCG	897
Qy	2308	CATCCCAGTCACGGCTGCAGATGATTCAGTTGAGGAAAGTGTCAAGCTGGAGTCCGGGCT	2367
Db	898	CATCCCAATGACGGCTGCAGATGATTCAGTTGAGGAAAGTGTCAAGCTGGAGTCCGGGCT	957
Qy	2368	TGTCAGGAGATGGCTTCTGTGCTCGGCACCAACCCACATCTGGTTGAGTTGGACCTGACA	2427
Db	958	TGTCAGGAGATGGCTTCTGTGCTTGGCACCAACCCACATCTGGTTGAGTTGGACCTGACA	1017
Qy	2428	GGAAATGCATCGAGGATTTGGGCTCAGGTTACTATGCCAGGACATGAGCCACCCAGTC	2487
Db	1018	GGAAATGCATCGAGGATTTGGGCTCAGGTTACTATGCCAGGACATGAGCCACCCAGTC	1077
Qy	2488	TGCAGACTACGGACTTTGTGGCTGAAGATCTGCCGCTCACTGCTGCTGCTGTGACGAG	2547
Db	1078	TGCAGACTACGGACTTTGTGGCTGAAGATCTGCCGCTCACTGCTGCTGCTGTGACGAG	1137
Qy	2548	CTGGCTCAACTCTCAGTGTGAACCAAGAGCTCGAGAGCTGAGACCTGAGCCTGGAATGAG	2607
Db	1138	CTGGCTCAACTCTCAGTGTGAACCAAGAGCTCGAGAGCTGAGACCTGAGCCTGGAATGAG	1197
Qy	2608	CTGGGGAACTCGGGGTGCTGCTGCTGTGTGAGGGCCTCAGGCATCCACGTCGCAAGCTC	2667
Db	1198	CTGGGGAACTCGGGGTGCTGCTGCTGTGTGAGGGCCTCAGGCATCCACGTCGCAAGCTC	1257
Qy	2668	CAGACCTCGGGTGGGCATCTGCCGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTG	2727
Db	1258	CAGACCTCGGGTGGGCATCTGCCGCTGGGCTCTGCCGCTGTGAGGCTCTTCTGTG	1317
Qy	2728	GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTTGATGTTTCAACGACCTGGGAGAC	2787
Db	1318	GTGCTCCAGGCCAACCAACACTCCGGGAGCTGGACTTTGATGTTTCAACGACCTGGGAGAC	1377
Qy	2788	TGGGGCTGTGTTGTGGCTGAGGGCTGCACAATCCCGCTCGACACTCCAGAACTG	2847
Db	1378	TGGGGCTGTGTTGTGGCTGAGGGCTGCACAATCCCGCTCGACACTCCAGAACTG	1437
Qy	2848	TGGCTGATAGCTGTGSCCTCAGAGCCAGGCTGTGAGAACTTTTACTTTCACCTGGGG	2907
Db	1438	TGGCTGATAGCTGTGSCCTCAGAGCCAGGCTGTGAGAACTTTTACTTTCACCTGGGG	1497
Qy	2908	ATCAACAGACCTTGACCGACTTTACTTGACCAACAAACGCCCTAGGGGACACAGGTGC	2967
Db	1498	ATCAACAGACCTTGACCGACTTTTACTTGACCAACAAACGCCCTAGGGGACACAGGTGC	1557
Qy	2968	CGACTCTTTGCAAGGGCTGAGGCATCTGGCTGCAAACTCCGAGTCTCTGGTTATTT	3027
Db	1558	CGACTCTTTGCAAGGGCTGAGGCATCTGGCTGCAAACTCCGAGTCTCTGGTTATTT	1617
Qy	3028	GGGATGGACTGAATAAATGACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAACCT	3087
Db	1618	GGGATGGACTGAATAAATGACCCACAGTAGGTTGGCAGCGCTTCGAGTAACAAACCT	1677
Qy	3088	TATTTGCAATTTGGCTGCTGA	3108
Db	1678	TATTTGCAATTTGGCTGCTGA	1698

RESULT 14

AX684291	AX684291	4931 bp	DNA	linear	PAT 29-MAR-2003
LOCUS	Sequence	1	from Patent	W002052011.	
DEFINITION	AX684291				
ACCESSION	AX684291				
VERSION	AX684291.1	GI:29371161			
KEYWORDS					
SOURCE	Human sapiens (human)				
ORGANISM	Human sapiens				
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Mammalia; Hominiidae; Homo.				
REFERENCE	1				
AUTHORS	Feder, J., Ramanathan, C. and Mintier, G.				
TITLE	Human leucine-rich repeat containing protein, hlrrb1, expressed				

Db 1974 AAATCTATTGAGGAGCAGGACCTCCGAGACACGGCTAGACGGGAGAGGCTCTCT 2033
QY 1396 GCCTTCTCAACATGAACATCTTCCAGAGGACATCAACTGTGAGAGG 1443
Db 2034 GCCTTCTCAACATGAACATCTTCCAGAGGACATCAACTGTGAGAGG 2081

RESULT 15
AF231021 2494 bp mRNA linear PRI 02-MAR-2001
LOCUS Homo sapiens leucine-rich-repeat protein RNO2 mRNA, complete cds.
DEFINITION
ACCESSION AF231021
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2494)
Shami,P.J., Kanai,N., Wang,L.Y., Vreeke,T.M. and Parker,C.J.
Identification and characterization of a novel gene that is
upregulated in leukemia cells by nitric oxide
Unpublished
2 (bases 1 to 2494)
Shami,P.J., Kanai,N., Wang,L.Y., Vreeke,T.M. and Parker,C.J.
Direct Submission
Submitted (02-FEB-2000) Medicine, University of Utah and SLC VA
Medical Centers, Box 151M, 500 Foothill Boulevard, Salt Lake City,
UT 84148, USA

FEATURES
Location/Qualifiers
1..2494
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1309..2172
/note="similar to the Homo sapiens ribonuclease/angiogenin
inhibitor; expressed in hematopoietic cells; upregulated
by nitric oxide in leukemia cells"
/codon_start=1
/product="leucine-rich-repeat protein RNO2"
/protein_id="AAK14942.1"
/db_xref="GI:13182797"
translation="MSQAWHTSVSPATQEAAGAGLQPRQRRLWKRCRTSSACED
LSAALIANKNLTRMDLNSGVGPGFMILLCEGLRHPQCRLOMIOIRKCOLSSGACQEM
ASVLGTHPLVELDITGNALDGLRLCCQGRHPVCKRTLWIKIKNLTPAACDELA
STLSVNSQLRELDLSINELGDLVLLCEGLRHPVCKRTLWIKIKNLTPAACDELYF
TLGINQTLDLYLTNNALGDTGVRLCKRLSHPGCKLRVLFGMDLNMKMTSHLAL
RVTKPYLDIGC"

ORIGIN
Query Match 16.6%; Score 515; DB 9; Length 2494;
Best Local Similarity 100.0%; Pred. No. 1.4e-277;
Matches 515; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2165 GCCTGAAGAGGTCCGCATCTCCAGCTCAGCTCGAGACCTCTCTGAGCTTCATAG 2224
Db 1400 GCCTGAAGAGGTCCGCATCTCCAGCTCAGCTCGAGACCTCTCTGAGCTTCATAG 1459
QY 2225 CCAATAGAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCAGGCATGA 2284
Db 1460 CCAATAGAATTTGACAGGATGGATCTCAGTGGCAACGGCGTTGGATTCCAGGCATGA 1519
QY 2285 TGCTGCTTTGAGGGCTCGCGGATCCCGAGTCGAGGTCGAGATGATTGATTGAGGA 2344
Db 1520 TGCTGCTTTGAGGGCTCGCGGATCCCGAGTCGAGGTCGAGATGATTGATTGAGGA 1579
QY 2345 AGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGCTCGGCACCAACCCAC 2404
Db 1580 AGTGTGAGTGGAGTCCGGGGCTTGTGAGGAGATGGCTTCTGCTCGGCACCAACCCAC 1639
QY 2405 ATCTGCTTGAAGTGGACCTGACAGGAATGACCTGAGGATTGGCGCTGAGGTTACTAT 2464

Db 1640 ATCTGTTGAGTTGGACCTGACAGGAAATGCACTGGAGGATTTGGGCTGAGGTTACTAT 1699
QY 2465 GCCAGGAGCTGAGGCAACCCAGTCTGACAGACTAGCGACTTTGTGGCTGAAGATCTGCCGCC 2524
Db 1700 GCCAGGAGCTGAGGCAACCCAGTCTGACAGACTAGCGACTTTGTGGCTGAAGATCTGCCGCC 1759
QY 2525 TCACTGCTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAG 2584
Db 1760 TCACTGCTGCTGCTGTGACGAGCTGGCCTCAACTCTCAGTGTGAACACAGAGCCTGAGAG 1819
QY 2585 AGCTGAGCCTGAGGCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTGTGTGAGGGCC 2644
Db 1820 AGCTGAGCCTGAGGCTGAATGAGCTGGGGACCTCGGGGTGCTGCTGCTGTGTGAGGGCC 1879
QY 2645 TCAGGCATCCACAGCTGCAAGCTCCAGACCTCGCGG 2679
Db 1880 TCAGGCATCCACAGCTGCAAGCTCCAGACCTCGCGG 1914

Search completed: August 8, 2004, 14:03:09
Job time : 11833 secs

Result No.	Query			Description		
	Score	Match	Length	DB	ID	
1	3108	100.0	3108	6	AD3636416	Human PAA
2	3057	98.4	3300	6	AAU47129	Pyrim dom
3	2475	79.6	3186	6	AAU44363	Human PYR
4	2096	67.4	3306	9	AD303016	Human nov
5	1356	43.6	2158	6	ABS78719	Human cDN
6	1212	39.0	1683	7	ACD03624	Novel hum
7	1116	35.9	1800	7	ACD03623	Novel hum
8	768	24.7	4931	6	ABL59333	Nucleotid
9	520	16.7	591	9	AD32201	Human nov
10	385	12.4	487	8	ACH36225	Human nov
11	346	11.1	479	3	AACT76566	Human end
12	289	9.3	594	4	AACT76566	Human ORF
13	289	9.3	3306	9	AD303016	Human nov
14	282	9.1	282	4	AAK50424	Human bon
15	254	8.2	506	5	AAK568757	Human bon
16	173	5.6	515	4	AAK37572	DNA encod
17	171	5.5	171	4	AAK50441	Human bon
18	152	4.9	289	6	AAK67173	Human imm
19	148	4.8	299	6	AAU44361	Human PYR
20	143	4.6	308	6	ABL59334	Expressed
21	136	4.4	352	4	AAK56010	Human imm
22	69	2.2	80	6	ABL59337	PCR prime
23	52	1.7	919	5	AAK568285	DNA encod

1501 CTGGACGAGGGAGGGGGGAGCCAGCCAGAGCCAGAGCCAGAGCTGTGACCGAG 1560
1501 CTGGACGAGGGAGGGGGGAGCCAGCCAGAGCCAGAGCCAGAGCTGTGACCGAG 1560
1561 TAGCGCTTTTCTGAAGAGCTTCTGCGCACTCACCAGCGCTTCTGTTGACCTCTG 1620
1561 TAGCGCTTTTCTGAAGAGCTTCTGCGCACTCACCAGCGCTTCTGTTGACCTCTG 1620
1621 AAGCGAGACACGAGGACCTTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGACATC 1680
1621 AAGCGAGACACGAGGACCTTGGAGAGAGTCTCTGCTGGAAGTCTCGCCGACATC 1680
1681 AAGTGGACCTGTGAGTGGATCCAAAGCAAAGCTCAGAGCGAGCTCCACCTGCG 1740
1681 AAGTGGACCTGTGAGTGGATCCAAAGCAAAGCTCAGAGCGAGCTCCACCTGCG 1740
1741 CAGGGCTCTTGGAGTCTTACGCTCTTGTAGAGATCCAGAGGAGGAGTTATCCAG 1800
1741 CAGGGCTCTTGGAGTCTTACGCTCTTGTAGAGATCCAGAGGAGGAGTTATCCAG 1800
1801 CAGGGCTCTTGGAGTCTTACGCTCTTGTAGAGATCCAGAGGAGGAGTTATCCAG 1860
1801 CAGGGCTCTTGGAGTCTTACGCTCTTGTAGAGATCCAGAGGAGGAGTTATCCAG 1860
1861 ATGCTCTCTGCTTCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
1861 ATGCTCTCTGCTTCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
1921 GGCACCTACAGCGCGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1921 GGCACCTACAGCGCGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
1981 TTGCTGAGCTCAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
1981 TTGCTGAGCTCAGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
2041 GCGGCTCTGACCAATCCAACTGATAGAGCTGTCTGTGAGAGGAGGAGGAGGAGG 2100
2041 GCGGCTCTGACCAATCCAACTGATAGAGCTGTCTGTGAGAGGAGGAGGAGGAGG 2100
2101 AGCGGGGGGCTGAGCTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2101 AGCGGGGGGCTGAGCTGTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
2161 CTGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
2161 CTGAGGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
2221 ATAGCAATAGAAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
2221 ATAGCAATAGAAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2280
2281 ATGATGCTGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
2281 ATGATGCTGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
2341 AGGAAGTCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
2341 AGGAAGTCTGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
2401 CCACATCTGTTGAGTGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
2401 CCACATCTGTTGAGTGGAGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2460
2461 CTATGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2461 CTATGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
2521 CGCCTCACTGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
2521 CGCCTCACTGCTGCTGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2580
2581 AGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2640

2581 AGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2640
2641 GGCCTCAGGATCCCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2700
2641 GGCCTCAGGATCCCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2700
2701 TCTGCCCTGAGGAGGAGCTTCTGCTGAGGAGGAGCTTCTGCTGAGGAGGAGCTT 2760
2701 TCTGCCCTGAGGAGGAGCTTCTGCTGAGGAGGAGCTTCTGCTGAGGAGGAGCTT 2760
2761 GACTTGAAGTTCAAGACCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2820
2761 GACTTGAAGTTCAAGACCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2820
2821 CATCCCGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2880
2821 CATCCCGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGG 2880
2881 TGTGAGAACTTTTACTTCAACCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 2940
2881 TGTGAGAACTTTTACTTCAACCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 2940
2941 AACAAAGCGCTTGGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 3000
2941 AACAAAGCGCTTGGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 3000
3001 TGCAGAACTCCGAGTCTCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 3060
3001 TGCAGAACTCCGAGTCTCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCT 3060
3061 TTGGCAGGCTTCCGAGTCAACAACTTATTTGGAGCTTGGCTGCTGCTGCTGCTG 3108
3061 TTGGCAGGCTTCCGAGTCAACAACTTATTTGGAGCTTGGCTGCTGCTGCTGCTG 3108

RESULT 2
AAL47129
ID AAL47129 standard; DNA; 3300 BP.
XX
AC AAL47129;
XX
DT 20-AUG-2002 (first entry)
XX
DE Pyrin domain containing protein NALP3/PY5-hs coding sequence.
XX
KW Pyrin domain; PYD domain; antiinflammatory; antiparkinsonian;
KW antiarteriosclerotic; antipsoriatic; antibacterial; virucide;
KW neuroprotective; antiarthritic; antirheumatic; antiasthmatic;
KW nephrotropic; osteopathic; nootropic; intracellular signal transduction;
KW inflammation; Alzheimer's disease; infection; peoriasis; asthma;
KW arteriosclerosis; multiple sclerosis; rheumatoid arthritis; sarcoidosis;
KW osteoarthritis; glomerulonephritis; gene; ds.
XX
OS Unidentified.
XX
PN WO200240668-A2.
XX
PD 23-MAY-2002.
XX
PF 30-OCT-2001; 2001WO-EP012545.
XX
PR 15-NOV-2000; 2000DE-01056687.
XX
PR 30-NOV-2000; 2000DE-01059595.
XX
PA (APOT-) APOTEC RES & DEV LTD.
XX
PI Tschopp J, Martinon F;
XX
XX WPI; 2002-427093/45.
DR P-PSDB; AAO17857.
XX
PT New DNA encoding protein with pyrin domain, useful for treating diseases

involving impaired signal transduction, particularly inflammation, also proteins and antibodies.

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

Claim 5; Fig 1; 116pp; German.

The present invention relates the DNA and their encoded proteins, where the proteins contain at least one PYD (pyrin) domain. These can be used to treat diseases associated with impaired intracellular signal transduction, particularly inflammation such as psoriasis, atherosclerosis, bacterial or viral infections (particularly meningitis and pneumonia), multiple sclerosis, rheumatoid arthritis, asthma, sarcoidosis, glomerulonephritis and osteoarthritis, and also Alzheimer's and Parkinson's diseases. The present sequence is a coding sequence of the invention.

Sequence 3300 BP: 726 A: 943 C: 979 G: 652 T: 0 U: 0 Other:

Query Match 98.4% Score 3057 DB 6 Length 3300;

| | | |
|-----------------------|--------|--------------|
| Query MACCM | 20.1% | 2003.03.01 |
| Best Local Similarity | 100.0% | Pred. No. 0: |

Best local similarity 100.0%, area: not set
Matches 3057: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

| | | | |
|----|-----|--|-----|
| Qy | 1 | ATGCTACGAACCGGACGAGCGGCGCTGTGTCCCTGTGCTCCACTACTTGGAGAACTC | 60 |
| | | | |
| Db | 1 | ATGCTACGAACCGGACGAGCGGCGCTGTGTCCCTGTGCTCCACTACTTGGAGAACTC | 60 |
| | | | |
| Qy | 61 | GAGGCTGTGAACTGGAAGAGTTCAAGTTATACCTGCGGACCGGACAGAGCTGGGAGAA | 120 |
| | | | |
| Db | 61 | GAGGCTGTGAACTGGAAGAGTTCAAGTTATACCTGCGGACCGGACAGAGCTGGGAGAA | 120 |
| | | | |
| Qy | 121 | GGCAAGATCCCTTGGGGAAGCATGGAGAAAGCCCGGTCCCTTGGAAATGGGCCAGCTGCTC | 180 |
| | | | |
| Db | 121 | GGCAAGATCCCTTGGGGAAGCATGGAGAAAGCCCGGTCCCTTGGAAATGGGCCAGCTGCTC | 180 |
| | | | |
| Qy | 181 | ATCACCCACTTTGGGGCCAGAGGAGGCTGTGAGGTTGGCTCTCAGCACCTTTTGAGGGGATA | 240 |
| | | | |
| Db | 181 | ATCACCCACTTTGGGGCCAGAGGAGGCTGTGAGGTTGGCTCTCAGCACCTTTTGAGGGGATA | 240 |
| | | | |
| Qy | 241 | AACAGGAGGACCTGTGGGAGAGAGGACAGAGAGGAGGACCTGTGTAGGGATCCCCAGGAA | 300 |
| | | | |
| Db | 241 | AACAGGAGGACCTGTGGGAGAGAGGACAGAGAGGAGGACCTGTGTAGGGATCCCCAGGAA | 300 |
| | | | |
| Qy | 301 | ACCTACAGGGACTATGTCCGACAGGAAATTCGGGCTCATGGAAGACCGCAATCGCGCCTA | 360 |
| | | | |
| Db | 301 | ACCTACAGGGACTATGTCCGACAGGAAATTCGGGCTCATGGAAGACCGCAATCGCGCCTA | 360 |
| | | | |
| Qy | 361 | GGGGAATGTGTCAACTCAGCCAACGGGTACACCGGGCTCTCTGTGTGAAGAGGACACTCA | 420 |
| | | | |
| Db | 361 | GGGGAATGTGTCAACTCAGCCAACGGGTACACCGGGCTCTCTGTGTGAAGAGGACACTCA | 420 |
| | | | |
| Qy | 421 | AACCCCATGAGGTCACGACGACGCTTCTGACACACAGCCCGGGACACGCGAGGACCGTG | 480 |
| | | | |
| Db | 421 | AACCCCATGAGGTCACGACGACGCTTCTGACACACAGCCCGGGACACGCGAGGACCGTG | 480 |
| | | | |
| Qy | 481 | GGACACAGGCTAGCCCCATCAAGATAGAGACCCCTCTTTAGGCCACAGAGAGCGCCCC | 540 |
| | | | |
| Db | 481 | GGACACAGGCTAGCCCCATCAAGATAGAGACCCCTCTTTAGGCCACAGAGAGCGCCCC | 540 |
| | | | |
| Qy | 541 | GAGCCACCGCGCACGTTGGTGTATGCAAGCGCGGAGGGATAGGCAAGTCAATGCTGGCA | 600 |
| | | | |
| Db | 541 | GAGCCACCGCGCACGTTGGTGTATGCAAGCGCGGAGGGATAGGCAAGTCAATGCTGGCA | 600 |
| | | | |
| Qy | 601 | CACAAGGTGATCTGGACTGGGGACGGGAGCTCTTCCAAAGGCAGATTTGATTATCTC | 660 |
| | | | |
| Db | 601 | CACAAGGTGATCTGGACTGGGGACGGGAGCTCTTCCAAAGGCAGATTTGATTATCTC | 660 |
| | | | |
| Qy | 661 | TTCTACATCAACTGCAGGAGATGAACACAGAGTGCACGGAAATGCAGCATGCAAGACCTC | 720 |
| | | | |
| Db | 661 | TTCTACATCAACTGCAGGAGATGAACACAGAGTGCACGGAAATGCAGCATGCAAGACCTC | 720 |
| | | | |
| Qy | 721 | ATCTTCAGCTGCTGCCTGAGCCGACGGCGCTCTCCAGGAGCTCATCGAGTTCCCGAG | 780 |
| | | | |
| Db | 721 | ATCTTCAGCTGCTGCCTGAGCCGACGGCGCTCTCCAGGAGCTCATCGAGTTCCCGAG | 780 |
| | | | |

1861 ATGGTCTCTCTGTTCTGTAAGGCGCTGCAGAGCGCCAGTGTGCACTGTATGTC 1920
1921 GCCACCTACAGCGCGGAGGAGACCGGAGGTGCTCGCAGGAGCGCACAGCTG 1980
1921 GCCACCTACAGCGCGGAGGAGACCGGAGGTGCTCGCAGGAGCGCACAGCTG 1980
1981 TTGGTGCAGCTCAGACCGAGAGGAGCGGTTCTGCTGGAGCGCTACAGTGAACATCTGGCA 2040
1981 TTGGTGCAGCTCAGACCGAGAGGAGCGGTTCTGCTGGAGCGCTACAGTGAACATCTGGCA 2040
2041 GCGGCCCTGTGCACCAATCCAACTGATAGAGTGTCTCTGACGAAATGCCCTGGGC 2100
2041 GCGGCCCTGTGCACCAATCCAACTGATAGAGTGTCTCTGACGAAATGCCCTGGGC 2100
2101 AGCCGGGGGTGAAGTGTCTGTAAGGACTCAGACACCCCACTGCAAACTTCAGAC 2160
2101 AGCCGGGGGTGAAGTGTCTGTAAGGACTCAGACACCCCACTGCAAACTTCAGAC 2160
2161 CTGAGGCTGAAGAGGTGCGGCATCTCCAGCTCAGCTGCGAGGAGCTCTCTCAGCTCTC 2220
2161 CTGAGGCTGAAGAGGTGCGGCATCTCCAGCTCAGCTGCGAGGAGCTCTCTCAGCTCTC 2220
2221 ATAGCCATAAGAAATTGACAGGATGATCTCAGTGGCAACGCGGTTGGATTCCAGGC 2280
2221 ATAGCCATAAGAAATTGACAGGATGATCTCAGTGGCAACGCGGTTGGATTCCAGGC 2280
2281 ATGATGCTGCTTTGCGAGGGCTCGGGCATCCCGAGTCCCGAGTGCAGATGATTCAGTTG 2340
2281 ATGATGCTGCTTTGCGAGGGCTCGGGCATCCCGAGTCCCGAGTGCAGATGATTCAGTTG 2340
2341 AGGAAGTGTGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2400
2341 AGGAAGTGTGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2400
2401 CCACATCTGTTGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2460
2401 CCACATCTGTTGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2460
2461 CTATGCCAGGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2520
2461 CTATGCCAGGAGTGTGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2520
2521 CGCCTCACTGCTGCTGCTGTCAGGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2580
2521 CGCCTCACTGCTGCTGCTGTCAGGAGTGTGCTGTCAGGAGATGGCTTCTGCTCGGCCCAAC 2580
2581 AGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2640
2581 AGAGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2640
2641 GGCCTCAGGCTATCCACCTGCAAGCTCCAGACCTCGGCTGGGATCTGCGGCTGGGC 2700
2641 GGCCTCAGGCTATCCACCTGCAAGCTCCAGACCTCGGCTGGGATCTGCGGCTGGGC 2700
2701 TCTGCGGCTGTGAGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
2701 TCTGCGGCTGTGAGGCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2760
2761 GACTTGAGTTTCAAGCACTGGAGACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
2761 GACTTGAGTTTCAAGCACTGGAGACTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
2821 CATCCGCTGCAGACTCCAGAACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
2821 CATCCGCTGCAGACTCCAGAACTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2880
2881 TGTGAGAACTTTTACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
2881 TGTGAGAACTTTTACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2940
2941 AACACGCCCTTAGGGGACACAGGTGTCGAGTGTCTTGTGAAAGCGGCTGAGCCATCCTGGC 3000

Db 2941 AACACGCCCTTAGGGGACACAGGTGTCGAGTGTCTTGTGAAAGCGGCTGAGCCATCCTGGC 3000
QY 3001 TCGAACTCCGAGTCTCTGTTATTTGGATGGACCTGAATAAAATGACCCACAGT 3057
Db 3001 TCGAACTCCGAGTCTCTGTTATTTGGATGGACCTGAATAAAATGACCCACAGT 3057
RESULT 3
AAL44363
ID AAL44363 standard; cDNA; 3186 BP.
XX
AC AAL44363;
XX
DT 31-OCT-2002 (first entry)
XX
DE Human PYRIN-8 cDNA sequence #2.
XX
KW Human; gene; ss; gene therapy; PYRIN; stress-related response;
KW apoptotic response; inflammatory disease; Crohn's disease; multiple sclerosis; cancer;
KW immune system disorder; Crohn's disease; arthritis; neurological disease;
KW leukaemia; autoimmune disorder; Crohn's disease; arthritis; neurological disease;
KW Alzheimer's disease; Parkinson's disease; chromosomal mapping;
KW tissue typing; forensic biology; predictive medicine; pharmacogenomics;
KW transcription profiling; PYRIN-8.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..3186
FT /tag= a
FT /product= "Human PYRIN-8 protein #2"
XX
XX WO200261049-A2.
XX
XX 08-AUG-2002.
XX
XX 31-JAN-2002; 2002WO-US002967.
XX
XX 31-JAN-2001; 2001US-0265231P.
XX 10-SEP-2001; 2001US-0318645P.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX (AMHP) WYETH.
XX
XX Bertin J, Wang W, Blatcher M;
XX
XX WPI; 2002-627477/67.
XX P-PSDB; AAO15590.
XX
XX New PYRIN polypeptides and nucleic acids useful for modulating and
XX diagnosing stress-related, apoptotic and inflammatory responses, or for
XX treating inflammatory and immune system disorders, cancers, or
XX neurological diseases.
XX
XX Claim 4; Fig 8; 167pp; English.
XX
XX The invention comprises the amino acid and coding sequences of human
XX PYRIN proteins. The PYRIN protein and DNA sequences of the invention are
XX useful for modulating and diagnosing stress-related, apoptotic and
XX inflammatory responses. The PYRIN protein and DNA sequences are useful
XX for treating inflammatory disorders and immune system disorders (e.g.
XX Crohn's disease, reactive arthritis, multiple sclerosis, contact
XX dermatitis, psoriasis, graft rejection, allergies, viral infections and
XX bacterial infections); cancer (e.g. leukaemia); autoimmune disorders
XX (e.g. systemic lupus erythematosus and arthritis); and neurological
XX diseases (e.g. Alzheimer's disease and Parkinson's disease). The PYRIN
XX protein and DNA sequences may also be used in screening assays, detection
XX assays (e.g. chromosomal mapping, tissue typing or forensic biology),
XX predictive medicine (e.g. diagnostic assays, clinical trials and
XX pharmacogenomics) and transcription profiling. The present DNA sequence
XX encodes a human PYRIN-8 protein
XX
XX Sequence 3186 BP; 701 A; 908 C; 945 G; 632 T; 0 U; 0 Other;

| | | | | | | | | | |
|---|------|--|------|----|------|--|------|--|--|
| Query Match 79.6%; Score 2475; DB 6; Length 3186;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2815; Conservative 0; Mismatches 2; Indels 3; Gaps 1; | | | | | | | | | |
| QY | 289 | GATCCCGAGAAACCTACAGGACTATGTCCGAGGAAATTCGGGTCTATGGAAGACCGC | 348 | QY | 1309 | TTGGCGGCGAGATGGGCTCTGGAAATCAGAAATCTATTATTGAGGAGCAGGACCTCCGGAAG | 1368 | | |
| DB | 370 | GATCCCGAGAAACCTACAGGACTATGTCCGAGGAAATTCGGGTCTATGGAAGACCGC | 429 | DB | 1390 | TTGGCGGCGAGATGGGCTCTGGAATCAGAAATCTTAITTTGAGGAGCAGGACCTCCGGAAG | 1449 | | |
| QY | 349 | AATGCGCGCTAGGGAATGTGTCAACCTCAGCCAGCGGTACACCGCGTCTCTGCTGGTG | 408 | QY | 1369 | CACGGCTTAGACGGGAGACGCTCTCTGCTTCCATAACATGAACATCTTCCAGAAGGAC | 1428 | | |
| DB | 430 | AATGCGCGCTTAGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGTCTCTGCTGGTG | 489 | DB | 1450 | CACGGCTTAGACGGGAGACGCTCTCTGCTTCCATAACATGAACATCTTCCAGAAGGAC | 1509 | | |
| QY | 409 | AAGGAGCACTCAAAACCCATGACAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC | 468 | QY | 1429 | ATCAACTGTGAGAGGTACTACAGCTTCACTCCTGAGTTTCCAGGAATTTCTTTGACGCT | 1488 | | |
| DB | 490 | AAGGAGCACTCAAAACCCATGACAGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACAC | 549 | DB | 1510 | ATCAACTGTGAGAGGTACTACAGCTTCACTCCTGAGTTTCCAGGAATTTCTTTGACGCT | 1569 | | |
| QY | 469 | GCAGAGCCGTGGGACACCAAGGTAGCCCATCAAGATAGAGACCTTCTTTGAGCCAGAC | 528 | QY | 1489 | ATGTAATATATCTCTGACGAGGGGAGGGCGGGGAGCCAGACACAGGACGTCGACCCAG | 1548 | | |
| DB | 550 | GCAGAGCCGTGGGACACCAAGGTAGCCCATCAAGATAGAGACCTTCTTTGAGCCAGAC | 609 | DB | 1570 | ATGTAATATATCTCTGACGAGGGGAGGGCGGGGAGCCAGACACAGGACGTCGACCCAG | 1629 | | |
| QY | 529 | GAGGAGCGCCGAGCCACCGCAGCCGTCATCAAGCGCGGAGGATAGGCAAG | 588 | QY | 1549 | CTGTGACCGAGTACGCGTTTTCTCAAAGGAGCTTCTGGCACTCACACGCGCTTCTCTG | 1608 | | |
| DB | 610 | GAGGAGCGCCGAGCCACCGCAGCCGTCATGCAAGCGCGGAGGATAGGCAAG | 669 | DB | 1630 | CTGTGACCGAGTACGCGTTTTCTGAAAGGAGCTTCTTGBCACTCACACGCGCTTCTCTG | 1689 | | |
| QY | 589 | TCCATGCTGGCACACAAGGTGATGCTGGAATGGGCGGACGGGAGCTTCTTCAAGGCAGA | 648 | QY | 1609 | TTTGGACTCTCTGAACGAGGAGACCAAGGAGCCACTTGGAAGAGTCTCTGCTGGAAGGTC | 1668 | | |
| DB | 670 | TCCATGCTGGCACACAAGGTGATGCTGGAATGGGCGGACGGGAGCTTCTTCAAGGCAGA | 729 | DB | 1690 | TTTGGACTCTCTGAACGAGGAGACCAAGGAGCCACTTGGAAGAGTCTCTGCTGGAAGGTC | 1749 | | |
| QY | 649 | TTTGATATCTCTTCTACATCAACTGACGAGGAGATGAACAGAGTGCACCGGAATGCAGC | 708 | QY | 1669 | TCGCGGCACATCAAGATGAGACCTGTTGCACTGGATCCAAAGCAAGCTCAGAGCGACGC | 1728 | | |
| DB | 730 | TTTGATATCTCTTCTACATCAACTGACGAGGAGATGAACAGAGTGCACCGGAATGCAGC | 789 | DB | 1750 | TCGCGGCACATCAAGATGAGACCTGTTGCACTGGATCCAAAGCAAGCTCAGAGCGACGC | 1809 | | |
| QY | 709 | ATGCAAGACCTCATCTTTCAGCTGCTGAGCCAGCGGCTCTCCAGGAGCTCATC | 768 | QY | 1729 | TCCACCTGACAGCGGCTCTTGAGTCTTTCAGCTGCTTCTACGAGTCCAGAGTCCAGAGGAG | 1788 | | |
| DB | 790 | ATGCAAGACCTCATCTTTCAGCTGCTGAGCCAGCGGCTCTCCAGGAGCTCATC | 849 | DB | 1810 | TCACACCTGACAGGAGCTTCTTGAGTCTTTCAGCTGCTTGTACGAGTCCAGAGGAG | 1869 | | |
| QY | 769 | CGAGTTTCCGAGCGCTCTCTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC | 828 | QY | 1789 | GAGTTTATCCAGCAGGCGCTTGAGCCACTTTCAGGTGATGCTGGTCAACAACTTGCTCTCC | 1848 | | |
| DB | 850 | CGAGTTTCCGAGCGCTCTCTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTC | 909 | DB | 1870 | GAGTTTATCCAGCAGGCGCTTGAGCCACTTTCAGGTGATGCTGGTCAACAACTTGCTCTCC | 1929 | | |
| QY | 829 | CACGATCCTCAGGAGCCCTGCTGCTGCTGGAGGAGAAACGGCCGACGGAGCTGCTT | 888 | QY | 1849 | AAGATGAGACACATGCTCTCTGTTCTGTGTGAAGCGCTGAGAGCGCCAGGTGCTG | 1908 | | |
| DB | 910 | CACGATCCTCAGGAGCCCTGCTGCTGCTGGAGGAGAAACGGCCGACGGAGCTGCTT | 969 | DB | 1930 | AAGATGAGACACATGCTCTCTGTTCTGTGTGAAGCGCTGAGAGCGCCAGGTGCTG | 1989 | | |
| QY | 889 | CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACACGG | 948 | QY | 1909 | CACCTGTATGCGCCACCTACAGCGCGAGCGGGAAGACCGCGAGGTGCTCCGACAGA | 1968 | | |
| DB | 970 | CTTAAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCAACACACGG | 1029 | DB | 1990 | CACCTGTATGCGCCACCTACAGCGCGAGCGGGAAGACCGCGAGGTGCTCCGACAGA | 2049 | | |
| QY | 949 | CCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCAACCCAGGCACTGTGGAGATCCTG | 1008 | QY | 1969 | GCGCACACGCTGTTGTTGTCAGCTCAGACACAGAGAGACCGTCTGTGTGACGCGCTACAGT | 2028 | | |
| DB | 1030 | CCACGGCTTTGGAGAGCTCCACCGTCTGCTGGAGCAACCCAGGCACTGTGGAGATCCTG | 1089 | DB | 2050 | GCGCACACGCTGTTGTTGTCAGCT--ACCAGAGAGACCGTCTGTGTGACGCGCTACAGT | 2106 | | |
| QY | 1009 | GGCTTCTCTGAGGAGAAAGAGGAATACTTCTCAAGTATTTCCAAATTCGAGAGCAG | 1068 | QY | 2029 | GAAATCTGGCAGCGGCGCTGTGCAACAACTGATAGAGTGTCTCTGTGTACCGA | 2088 | | |
| DB | 1090 | GGCTTCTCTGAGGAGAAAGAGGAATACTTCTCAAGTATTTCCAAATTCGAGAGCAG | 1149 | DB | 2107 | GAAATCTGGCAGCGGCGCTGTGCAACAACTGATAGAGTGTCTCTGTGTACCGA | 2166 | | |
| QY | 1069 | GCGGGCCAACTTCAATTAAGTGAAGGACCAAGGCTCTTTCACCATGCTGCTTGGTC | 1128 | QY | 2089 | AATGCCCTGGCAGCGCGGGGTGAAGCTGCTGTGTCAAGGACTCAGACACCCCAACTGC | 2148 | | |
| DB | 1150 | GCGGGCCAACTTCAATTAAGTGAAGGACCAAGGCTCTTTCACCATGCTGCTTGGTC | 1209 | DB | 2167 | AATGCCCTGGCAGCGCGGGGTGAAGCTGCTGTGTCAAGGACTCAGACACCCCAACTGC | 2226 | | |
| QY | 1129 | CCCCGTGTGCTGGGTGTGTGACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG | 1188 | QY | 2149 | AAACTTTCAGAACTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTGCGAGGACCTC | 2208 | | |
| DB | 1210 | CCCCGTGTGCTGGGTGTGTGACCTGCTCCAGCAGCAGCTGGAGGGTGGGGGGCTG | 1269 | DB | 2227 | AAACTTTCAGAACTGAGGCTGAAGAGGTGCGCATCTCCAGCTCAGCTGCGAGGACCTC | 2286 | | |
| QY | 1189 | TTGAGACAGAGCTCCAGGACACCACTGCACTGTACATGCTCTACCTGCTGAGTCTGATG | 1248 | QY | 2209 | TCTCAGCTCTCATAGCCCAATGAAGATTTTGACAAGGATGGATCTCAGTGGCAACGGGCTT | 2268 | | |
| DB | 1270 | TTGAGACAGAGCTCCAGGACACCACTGCACTGTACATGCTCTACCTGCTGAGTCTGATG | 1329 | DB | 2287 | TCTCAGCTCTCATAGCCCAATGAAGATTTTGACAAGGATGGATCTCAGTGGCAACGGGCTT | 2346 | | |
| QY | 1249 | CAACCCAGCGGGGGCGCGCTCCAGCGCCCAACCAACAGAGAGGGTTGTGCTCC | 1308 | QY | 2269 | GGATTTCCAGGCATGATGCTGCTTTTGGAGGGGCTTCGGGCATCCCACTGAGGCTGCGAG | 2328 | | |
| DB | 1330 | CAACCCAGCGGGGGCGCGCTCCAGCGCCCAACCAACAGAGAGGGTTGTGCTCC | 1389 | DB | 2347 | GGATTTCCAGGCATGATGCTGCTTTTGGAGGGGCTTCGGGCATCCCACTGAGGCTGCGAG | 2406 | | |
| | | | | QY | 2329 | ATGATTCAGTTGAGGAAGTGTGAGTGGGTTCGGGGCTTGTGAGGAGATGGCTTCTGTG | 2388 | | |
| | | | | DB | 2407 | ATGATTCAGTTGAGGAAGTGTGAGTGGGTTCGGGGCTTGTGAGGAGATGGCTTCTGTG | 2466 | | |
| | | | | QY | 2389 | CTCGCCACCAACCCCATCTGTTGAGTTGAGCTTGACAGGAAATGCATCGAGGAGATTG | 2448 | | |

| | | | |
|----|------|---|------|
| Db | 2467 | CTTGGCACCAACCCACATCTGGTTGAGTTGGACTGGACAGAAATGCACCTGGAGGATTTG | 2526 |
| Qy | 2449 | GGCCTGAGGTTACTATGCCAGGGACTGAGGCACCCAGTCTGCAGACTACGGACTTTGTGG | 2508 |
| Db | 2527 | GGCCTGAGGTTACTATGCCAGGGACTGAGGCACCCAGTCTGCAGACTACGGACTTTGTGG | 2586 |
| Qy | 2509 | CTGAAGATCTCCCGCCTCACTGCTGCTGTGTGACGAGCTGGCCTCAACTCTCAFGTGTG | 2568 |
| Db | 2587 | CTGAAGATCTCCCGCCTCACTGCTGCTGTGTGACGAGCTGGCCTCAACTCTCAFGTGTG | 2646 |
| Qy | 2569 | AACACAGAGCTGAGAGAGCTGGACCTCAGAGCTGAAATGAGCTGGGGACCTCGGGGTGCTG | 2628 |
| Db | 2647 | AACACAGAGCTGAGAGAGCTGGACCTGAGCTGAAATGAGCTGGGGACCTCGGGGTGCTG | 2706 |
| Qy | 2629 | CTGCTGTGTGAGGGCTCTCAGGCATCCACGTGCAAGCTCCAGAGCTTCGCGTTGGGCATC | 2688 |
| Db | 2707 | CTGCTGTGTGAGGGCTCTCAGGCATCCACGTGCAAGCTCCAGAGCTTCGCGTTGGGCATC | 2766 |
| Qy | 2689 | TGCGCGCTGGCTCTGCGCGCTGTGAGGTCCTTTCTGTGTGTCTCAGGCCAACCAAC | 2748 |
| Db | 2767 | TGCGCGCTGGCTCTGCGCGCTGTGAGGTCCTTTCTGTGTGTCTCAGGCCAACCAAC | 2826 |
| Qy | 2749 | CTCGGAGCTGGACTTGAGTTTCAACGACTGGGAGCTGGGCTGTGCTGTGCTGGCT | 2808 |
| Db | 2827 | CTCGGAGCTGGACTTGAGTTTCAACGACTGGGAGCTGGGCTGTGCTGTGCTGGCT | 2886 |
| Qy | 2809 | GAGGGGTGCAACATCCCGCTCGAGACTCCAGAACTGTGGCTGGATAGCTGTGGCCTC | 2868 |
| Db | 2887 | GAGGGGTGCAACATCCCGCTCGAGACTCCAGAACTGTGGCTGGATAGCTGTGGCCTC | 2946 |
| Qy | 2869 | ACAGCCAGGCTTGTGAGATCTTTACTTACCTGGGGATCAACAGACCTTGAACCGAC | 2928 |
| Db | 2947 | ACAGCCAGGCTTGTGAGATCTTTACTTACCTGGGGATCAACAGACCTTGAACCGAC | 3006 |
| Qy | 2929 | CTTTACTGACCAACAAAGCCCTAGGGGACACAGGTGTCCGATGCTTTTGCAGAGCGCTG | 2988 |
| Db | 3007 | CTTTACTGACCAACAAAGCCCTAGGGGACACAGGTGTCCGATGCTTTTGCAGAGCGCTG | 3066 |
| Qy | 2989 | AGCCATCTGGCTGCAAACTCCGAGTCTCTGTGTTATTTGGGATGGACCTGAAATAAATG | 3048 |
| Db | 3067 | AGCCATCTGGCTGCAAACTCCGAGTCTCTGTGTTATTTGGGATGGACCTGAAATAAATG | 3126 |
| Qy | 3049 | ACCCAGTAGGTTGCGAGGCTTCAGTAACAAACCTTATTTGGACATGCGTCTGCTGA | 3108 |
| Db | 3127 | ACCCAGTAGGTTGCGAGGCTTCAGTAACAAACCTTATTTGGACATGCGTCTGCTGA | 3186 |

RESULT, T 4

RESULT 4
ADC30316

ID ADC30316 standard; cDNA; 3306 BP.

AC ADC30316:

DT 18-DEC-2003 (first entry)

XX DE Human novel cDNA sequence. SEQ ID NO: 398

Human; diagnostic; drug screening; forensics; gene mapping;
 biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 ulcers; osteoporosis; autoimmune disease; cancer;
 molecular weight marker; food supplement; antiparkinsonian; neurotropic;
 neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;
 antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 gene therapy; chromosome 19; gene; ss.

OS Homo sapiens.

XX PN WO2003029271-A2

XX PD 10-APR-2003

XX

24-SEP-2002; 2002WO-US030474.

24-BEF-200Z; 200ZWO-US030474.

DT007 ; T007-JES-47

(HYSE-) HYSEQ INC.

Tang TY, Zhang J, Ren F, Xue AJ, Zhao OA, Wang J, Wehrma

Tang Y, Zhang J, Ren F, Zhou P, Ghosh M, Wang D, M

Haley-Vicente D, BR

WPI; 2003-371981/

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 1. SEQ ID NO 398. 1195bp. English

The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 9% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 contig sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC3262) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published/pct](http://wipo.int/pub/published/pct) sequences.

Sequence 3306 BP: 718 A: 958 C: 939 G: 691 T: 0 U: 0 Other: .

Query Match

| | |
|-----------------------|---------------------------------------|
| Query match | 67.4%; Score 2096; DB 9; Length 3306; |
| Best Local Similarity | 99.8%; Pred. No. 0; |

Total local similarity: 55.6%;
 Matches 2386; Conservative 0; Mismatched
 NO: 0; Fixed: 0

290 ATCCCCAGGAACCTACAGGGACTATGTCCGCAGGAAATTCCGGCTCATGGAAGACCGCA 349

230 ATCCCCAGGAAACCTACAGGGACATATGTCCTCCAGGAATATCCGGCTCATGGGAAGACCGCA 349

4398 ATCCCCAGGAAACCAACAGGGACATATGTCCGCAGGAAATTCGGCTCATGGAAGACCGCA 557

350 ATGCGCGCCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCCGGCTCCTGCTGGTGA 409

b
558 ATGCGCGCTAGGGGAATGTGTCAACCTCAGCCACCGGTACACCGGTCCTGCTGTGA 617

558 A T G C C G C C C T A G G G A A T G T G T C A A C C T C A G C C A C C G G T A C A C C G G C T C C T G C T G G T G A 617

410 AGGAGCACCTCAAACCCCATGCAGGTCCAGCAGCAGCTTCTGGACACAGGCCGGGACACG 469

Db 678 CGAGGACCGTGGACACCGCTAGGCCCAACAAGATAGAGACCCCTCTTTGAGCCGAGCG 737
Qy 530 AGCAGCGCCCCGAGCCACCCGCGACCCGCTGTGTCATGCAAGCGCGGACGAGGATAGGCAAGT 589
Db 738 AGAGCGCCCCGAGCCACCCGCGACCCGCTGTGTCATGCAAGCGCGGACGAGGATAGGCAAGT 797
Qy 590 CCAATGCTGGCACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAGGCAGAT 649
Db 798 CCAATGCTGGCACACAAGGTGATGCTGGACTGGGCGGACGGGAAGCTCTTCCAAGGCAGAT 857
Qy 650 TTGATATCTCTTCTTCAATCAATGACGAGGAGATGAACACAGAGTGCCACGGATGACGCA 709
Db 858 TTGATATCTCTTCTTCAATCAATGACGAGGAGATGAACACAGAGTGCCACGGATGACGCA 917
Qy 710 TGCAAGACCTCATCTTCAGCTGCTGGCTGAGCCAGCGGCCCTCTCCAGAGGCTCATCC 769
Db 918 TGCAGACCTCATCTTCAGCTGCTGGCTGAGCCAGCGGCCCTCTCCAGAGGCTCATCC 977
Qy 770 GAGTCCCGAGCGCCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTCC 829
Db 978 GAGTCCCGAGCGCCTCTCTTTTCATCATCGACGGCTTCGATGAGCTCAAGCCTTCTTTCC 1037
Qy 830 ACGATCCTCAGGACCTGCTGGCTCTGCTGGAGGAGAAACGSCCAGGAGCTGCTTC 889
Db 1038 ACGATCCTCAGGACCTGCTGGCTCTGCTGGAGGAGAAACGSCCAGGAGCTGCTTC 1097
Qy 890 TTAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGGC 949
Db 1098 TTAACAGCTTAATTCGGAAGAGCTGCTCCCTGAGCTATCTTTGCTCATCACACACGGC 1157
Qy 950 CCAGCGCTTTGGAGAAAGCTCCACCGTCTGCTGAGAGACCCACAGGATGAGAGATCCTGG 1009
Db 1158 CCAGCGCTTTGGAGAAAGCTCCACCGTCTGCTGAGAGACCCACAGGATGAGAGATCCTGG 1217
Qy 1010 GCTTCTCTGAGGACGAGGAGGAGAACTCTTCAAGTATTTCCAAATGCAAGAGGAGG 1069
Db 1218 GCTTCTCTGAGGACGAGGAGGAGAACTCTTCAAGTATTTCCAAATGCAAGAGGAGG 1277
Qy 1070 CGGCGCAAGCTTTCAAATTAACGCTGAGGAGCAACGAGCGCTCTTCCACCATGCTGCTGCC 1129
Db 1278 CGGCGCAAGCTTTCAAATTAACGCTGAGGAGCAACGAGCGCTCTTCCACCATGCTGCTGCC 1337
Qy 1130 CCTGTGTGCTGGGTGTGTGTAACCTGCTGACGAGACAGCTGAGAGGTGGGGGCTGT 1189
Db 1338 CCTGTGTGCTGGGTGTGTGTAACCTGCTGACGAGACAGCTGAGAGGTGGGGGCTGT 1397
Qy 1190 TGAGACAGAGCTCCAGGACCAACCTGCTGAGTACATGCTCTACCTGCTGAGTCTGATGC 1249
Db 1398 TGAGACAGAGCTCCAGGACCAACCTGCTGAGTACATGCTCTACCTGCTGAGTCTGATGC 1457
Qy 1250 AACCACAGCGGGGGCCCCGGCGCTCCAGCGCCCCACCCCAACACAGAGAGGTTGTGCTCCT 1309
Db 1458 AACCACAGCGGGGGCCCCGGCGCTCCAGCGCCCCACCCCAACACAGAGAGGTTGTGCTCCT 1517
Qy 1310 TGGCGGAGATGGGCTGTGGAATCAGAAATCCTATTGAGGAGGAGGACCTCCGGAAGC 1369
Db 1518 TGGCGGAGATGGGCTGTGGAATCAGAAATCCTATTGAGGAGGAGGACCTCCGGAAGC 1577
Qy 1370 ACGGCTAGACGGGGAAGAGCTCTCTGCTTCTCAACATGAACATCTTCAGAAAGGACA 1429
Db 1578 ACGGCTAGACGGGGAAGAGCTCTCTGCTTCTCAACATGAACATCTTCAGAAAGGACA 1637
Qy 1430 TCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAATTCAGGAAATCTTTGCAGCTA 1489
Db 1638 TCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAATTCAGGAAATCTTTGCAGCTA 1697
Qy 1490 TGTAATATATCTTGACGAGGAGGAGGCGGGGAGGCGCCAGACCGAGACCTGACCCAGGC 1549
Db 1698 TGTAATATATCTTGACGAGGAGGAGGCGGGGAGGCGCCAGACCGAGACCTGACCCAGGC 1757
Qy 1550 TGTTGACCGAGTACCGGTTTTCTGAAAGAGCTTCTCTGGCACTCACGAGCGCTTCCTGT 1609
Db 1758 TGTTGACCGAGTACCGGTTTTCTGAAAGAGGCTTCTCTGGCACTCACGAGCGCTTCCTGT 1817

Qy 1610 TTGGACTCTGTAAACGAGGAGACGAGGAGCCACCTGGAGAAAGTCTTCTGCTGGAGGTCT 1669
Db 1818 TTGGACTCTGTAAACGAGGAGACGAGGAGCCACCTGGAGAAAGTCTTCTGCTGGAGGTCT 1877
Qy 1670 CGCGGCACATCAAGATGGACCTGTCAGTGGATCCAAAGCAAGCTCAGAGCGACCGCT 1729
Db 1878 CGCGGCACATCAAGATGGACCTGTCAGTGGATCCAAAGCAAGCTCAGAGCGACCGCT 1937
Qy 1730 CCACCTGACGAGGGCTCTTGGAGTCTTTCAGTCTGTTGTACGAGATCCAGAGGAGG 1789
Db 1938 CCACCTGACGAGGGCTCTTGGAGTCTTTCAGTCTGTTGTACGAGATCCAGAGGAGG 1997
Qy 1790 AGTTATCCAGAGCGCCTCAGCOACTTCAGGFGATCGTGGTCAGCAACATTCCTTCCA 1849
Db 1998 AGTTATCCAGAGCGCCTCAGCOACTTCAGGFGATCGTGGTCAGCAACATTCCTTCCA 2057
Qy 1850 AGATGGAGCATGCTCTCTCTGTTCTGTGTGAAGCGCTGACGAGCGCCAGGTGCTGC 1909
Db 2058 AGATGGAGCATGCTCTCTCTGTTCTGTGTGAAGCGCTGACGAGCGCCAGGTGCTGC 2117
Qy 1910 ACTTGTATGGCGCCACCTCAGACGCGGAGCGGGAGAACCGCGAGGTGCTCCGACGAG 1969
Db 2118 ACTTGTATGGCGCCACCTCAGACGCGGAGCGGGAGAACCGCGAGGTGCTCCGACGAG 2177
Qy 1970 CGCACACGCTGTTGGTGCAGCTCAGACGAGAGGACCGTTCCTGCTGGAACGCTTACAGTG 2029
Db 2178 CGCACACGCTGTTGGTGCAGCT---ACGAGAGAGGACCGTTCCTGCTGGAACGCTTACAGTG 2234
Qy 2030 AACATCTGGCAGCGCCCTGTCACCAATCCAACTGTAGAGCTGTCCTGTACCGAA 2089
Db 2235 AACATCTGGCAGCGCCCTGTCACCAATCCAACTGTAGAGCTGTCCTGTACCGAA 2294
Qy 2090 ATGCCCTGGGCGAGCGGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCACTGCA 2149
Db 2295 ATGCCCTGGGCGAGCGGGGGTGAAGCTGCTCTCAAGGACTCAGACACCCCACTGCA 2354
Qy 2150 AACTTCAGAACTCAGGCTGAAGGCTGCGCATCTCCAGCTCAGCTGCGAGGACCTCT 2209
Db 2355 AACTTCAGAACTCAGGCTGAAGGCTGCGCATCTCCAGCTCAGCTGCGAGGACCTCT 2414
Qy 2210 CTGCAGCTCTCATAGCCCAATAAGAAATTTGACAGAGTGGATCTCAGTGGCAACGCGTTG 2269
Db 2415 CTGCAGCTCTCATAGCCCAATAAGAAATTTGACAGAGTGGATCTCAGTGGCAACGCGTTG 2474
Qy 2270 GATTCCAGGACATGATGCTGCTTTTCGAGGGCTCGCGCATCCCAGTGCAGGTGCAGA 2329
Db 2475 GATTCCAGGACATGATGCTGCTTTTCGAGGGCTCGCGCATCCCAGTGCAGGTGCAGA 2534
Qy 2330 TGATTCTAGTTGAGGAAAGTGTGAGTCCGGGGCTTGTTCAGGAGATGGCTTCTGTGC 2389
Db 2535 TGATTCTAGTTGAGGAAAGTGTGAGTCCGGGGCTTGTTCAGGAGATGGCTTCTGTGC 2594
Qy 2390 TCGGCACCAACCCACATCTGGTTCAGTTGACCTGACAGGAAATGCACTGGAGATTTGG 2449
Db 2595 TCGGCACCAACCCACATCTGGTTCAGTTGACCTGACAGGAAATGCACTGGAGATTTGG 2654
Qy 2450 GCCTGAGGTTACTATGCCAGGACTGAGGCAACCAAGTTCAGAGTACGAGCTTTGTGC 2509
Db 2655 GCCTGAGGTTACTATGCCAGGACTGAGGCAACCAAGTTCAGAGTACGAGCTTTGTGC 2714
Qy 2510 TGAAGATCTGCGGCTCATCTGCTGCTGTGAGAGCTGGCTCAACTCTCAGTGTGA 2569
Db 2715 TGAAGATCTGCGGCTCATCTGCTGCTGTGAGAGCTGGCTCAACTCTCAGTGTGA 2774
Qy 2570 ACCAGAGCTCAGAGAGCTGAGCTTGAATGAGCTGGGGACCTCGGGGTGCTGC 2629
Db 2775 ACCAGAGCTCAGAGAGCTGAGCTTGAATGAGCTGGGGACCTCGGGGTGCTGC 2834
Qy 2630 TSGCTGTGAGGGGCTCAGGCAATCCAGTGCAGAGCTCCAGACCTCGGG 2679
Db 2835 TSGCTGTGAGGGGCTCAGGCAATCCAGTGCAGAGCTCCAGACCTCGGG 2884

RESULT 5

ABS/78719
ID ABS78719 standard; cDNA; 2158 BP.
AC ABS78719;
XX
DT 16-DEC-2002 (first entry)
XX
DE Human cDNA encoding NAAP9, from INCYTE no.429930CB1.
XX
KW Human; ss; gene; nucleic acid associated protein; NAAP; cancer;
cell proliferative disease; cancer; atherosclerosis; hepatitis;
neurological disorder; Parkinson's disease; Alzheimer's disease; stroke;
epilepsy; developmental disorder; renal tubular acidosis; anaemia;
glaucoma; hypothyroidism; autoimmune disorder; AIDS;
inflammatory disorder; acquired immunodeficiency syndrome; allergy;
atopic dermatitis; arthritis; bacterial infection; viral infection;
parasitic infection; protozoal infection; fungal infection.
XX
OS Homo sapiens.
XX
PN WO200272630-A2.
XX
PD 19-SEP-2002.
XX
PF 07-FEB-2002; 2002WO-US003844.
XX
PR 09-FEB-2001; 2001US-0269118P.
PR 21-FEB-2001; 2001US-0270963P.
PR 22-FEB-2001; 2001US-0270858P.
PR 23-FEB-2001; 2001US-0271194P.
PR 07-MAR-2001; 2001US-0274071P.
PR 12-APR-2001; 2001US-0283496P.
PR 09-NOV-2001; 2001US-0344650P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Thornton M, Hafalia AJA, Lu DAM, Arvizu C, Swarnakar A, Lu Y;
PI Warren BA, Baughn MR, Tang YT, Lee EA, Yao MG, Ramkumar J, Khan FA;
PI Gandhi AR, Ding L, Yue H, Gietzen KJ, Walia NK, Thangavelu K;
PI Elliott VS, Marquis JP;
XX
DR WPI: 2002-723320/78.
DR P-PSDB; ABG97475.
XX
PT New human nucleic acid-associated proteins (NAAP), useful for diagnosing,
PT treating and preventing diseases or conditions associated with the
PT aberrant NAAP expression, e.g. cancer, hepatitis, AIDS, atherosclerosis,
PT infections.
XX
PS Claim 5; Page 160; 162pp; English.
XX
CC The invention relates to an isolated polypeptide comprising one of 10
CC human nucleic acid associated protein (NAAP1-10), or a biologically
CC active or immunogenic fragment of the polypeptide, and their encoding
CC nucleic acid. Also included are a recombinant polynucleotide comprising a
CC promoter sequence operably linked to the polynucleotide, a cell
CC transformed with the recombinant polynucleotide, a transgenic organism
CC comprising the recombinant polynucleotide, an anti-NAAP antibody,
CC screening for a compound that is effective as an antagonist or modulator
CC of NAAP, generating an expression profile of a sample containing the
CC polynucleotides and an array comprising different nucleotide molecules
CC affixed on a solid substrate, nucleotide molecule comprises a first
CC oligonucleotide or polynucleotide sequence specifically hybridisable with
CC at least 30 contiguous nucleotides of the target (NAAP) polynucleotide.
CC The polypeptides and polynucleotides are useful in diagnosing, treating
CC and preventing diseases or conditions associated with the decreased
CC expression or overexpression of NAAP, such as cell proliferative diseases
CC (e.g. cancer, atherosclerosis, hepatitis), neurological disorders
CC (Parkinson's disease, Alzheimer's disease, stroke, epilepsy),
CC developmental disorders (renal tubular acidosis, anaemia, glaucoma,
CC hypothyroidism), autoimmune/inflammatory disorders (AIDS (acquired

CC immunodeficiency syndrome), allergies, atopic dermatitis, arthritis) and
CC infections (e.g. bacterial, viral, parasitic, protozoal, fungal) and many
CC other diseases and disorders listed in the specification. These are also
CC useful in assessing the effects of exogenous compounds on the expression
CC of nucleic acid and amino acid sequences of NAAP. The NAAP or its
CC fragments are useful in screening compounds for effectiveness as agonist
CC or antagonist of the polypeptides, or in altering the expression of the
CC target polynucleotide and compounds that specifically bind to or modulate
CC the activity of the polypeptide. The microarray is useful in monitoring
CC or measuring protein-protein interactions, drug-target interactions, and
CC gene expression profiles. The present sequence encodes an NAAP protein
XX
SQ Sequence 2158 BP; 489 A; 578 C; 639 G; 452 T; 0 U; 0 Other;
Query Match 43.6%; Score 1356; DB 6; Length 2158;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 2; Indels 3; Gaps 1;
QY 1408 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 1467
DB 1 ATGAACATCTTCCAGAGGACATCAACTGTGAGAGGTACTACAGCTTCATCCACTTGAGT 60
QY 1468 TTCAGGAAATCTTTGACGCTATGATCTATATCTTGGACGAGGGGGGGGAGGC 1527
DB 61 TTCCAGGAAATCTTTGACGCTATGATCTATATCTTGGACGAGGGGGGGGAGGC 120
QY 1528 CCAGACGAGGACGTGACCCAGGCTGTTGACCGAGTACGCGTTTCTGAAAGAGGCTTCCTG 1587
DB 121 CCAGACGAGGACGTGACCCAGGCTGTTGACCGAGTACGCGTTTCTGAAAGAGGCTTCCTG 180
QY 1588 GCACCTCACCCAGCGCTTCTCTGTTGGACTCTCTGAACAGAGGAGACAGGAGCCACCTGGAG 1647
DB 181 GCACCTCACCCAGCGCTTCTCTGTTGGACTCTCTGAACAGAGGAGACAGGAGCCACCTGGAG 240
QY 1648 AAGAGTCTCTGCTGGAAAGGTCTGCGCCGACATCAAGATGGACCTTTGACGFGATCCAA 1707
DB 241 AAGAGTCTCTGCTGGAAAGGTCTGCGCCGACATCAAGATGGACCTTTGACGFGATCCAA 300
QY 1708 AGCAAGCTCAGAGGAGCGGCTCCACCTGACGAGGGGCTCTTGAGTCTTCTAGCTGC 1767
DB 301 AGCAAGCTCAGAGGAGCGGCTCCACCTGACGAGGGGCTCTTGAGTCTTCTAGCTGC 360
QY 1768 TTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTTCCAGGAGTTCACGAGTGATC 1827
DB 361 TTGTACGAGATCCAGGAGGAGGAGTTTATCCAGCAGGCGCTTCCAGGAGTTCACGAGTGATC 420
QY 1828 GTGTGACGAAATTCGCTCCAAAGATGGAGCACAATGTTCTCTCGTCTGTCTGAAGCGC 1887
DB 421 GTGTGACGAAATTCGCTCCAAAGATGGAGCACAATGTTCTCTCGTCTGTCTGAAGCGC 480
QY 1888 TGCAGGAGCGCCAGGCTGCTGACCTTGTATGGCGCCACTCAGCGCGGAGCGGGAAGAC 1947
DB 481 TGCAGGAGCGCCAGGCTGCTGACCTTGTATGGCGCCACTCAGCGCGGAGCGGGAAGAC 540
QY 1948 CGCGGAGGCTGCTCCGAGGAGCGCACAACGCTGTTGGTGACGCTCAGACCCAGAGAGGACC 2007
DB 541 CGCGGAGGCTGCTCCGAGGAGCGCACAACGCTGTTGGTGACGCT---ACCAGAGAGGACC 597
QY 2008 GTTCTGTGGACGCTACAGTGAACATCTGCGAGCGGCCCTGTGACCAATCCAAACCTG 2067
DB 598 GTTCTGTGGACGCTACAGTGAACATCTGCGAGCGGCCCTGTGACCAATCCAAACCTG 657
QY 2068 ATAGAGTGTCTGTGTACCGAAATTCGCTGGGAGCGGGGGGTGAAGCTGCTCTGTCAA 2127
DB 658 ATAGAGTGTCTGTGTACCGAAATTCGCTGGGAGCGGGGGGTGAAGCTGCTCTGTCAA 717
QY 2128 GGACTTCAGACACCCCACTGCAAACTTCAGAACCTTCCAGAGGTGAAGAGTGGCGATCTCC 2187
DB 718 GGACTTCAGACACCCCACTGCAAACTTCCAGAACCTTCCAGAGGTGAAGAGTGGCGATCTCC 777
QY 2188 AGCTCAGGCTCGAGGAGGACCTCTCTGACGCTCTCATAGCAATAGAAATTTGACAGGATG 2247
DB 778 AGCTCAGGCTCGAGGAGGACCTCTCTGACGCTCTCATAGCAATAGAAATTTGACAGGATG 837

| | | | |
|----|------|--|------|
| QY | 2248 | GATCTCAGTGGCAACGGCGTTTGAAATCCCAAGGCAATGAATGCTGCTTTGCGAGGGCCTGCGG | 2307 |
| Db | 838 | GATCTCAGTGGCAACGGCGTTTGAAATCCCAAGGCAATGAATGCTGCTTTGCGAGGGCCTGCGG | 897 |
| QY | 2308 | CATCCCCAGTGCAGGCTGCAGATGATTTCAGTTCAGGAAGTGCAGCTGGAAGTCCGGGGCT | 2367 |
| Db | 898 | CATCCCAATGCAGCTGCAGATGATTTCAGTTCAGGAAGTGCAGCTGGAAGTCCGGGGCT | 957 |
| QY | 2368 | TGTCAGGAGATGGCTTCTGTGCTCGGCACCAACCACATCTGGTTGAGTTGGACCTGACA | 2427 |
| Db | 958 | TGTCAGGAGATGGCTTCTGTGCTTGGCACCAACCCACATCTGGTTGAGTTGGACCTGACA | 1017 |
| QY | 2428 | GGAAATGCACGTGAGGATTTTGGGCTGAGGTTACTATGCCAGGGACTGAGGCAACCAATC | 2487 |
| Db | 1018 | GGAAATGCACGTGAGGATTTTGGGCTGAGGTTACTATGCCAGGGACTGAGGCAACCAATC | 1077 |
| QY | 2488 | TGCAGACTACGGACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCTGTGACGAG | 2547 |
| Db | 1078 | TGCAGACTACGGACTTTGTGGCTGAAGATCTGCGGCTCACTGCTGCTGCTGTGACGAG | 1137 |
| QY | 2548 | CTGGCTCTAACTCTCAGTGTGAACCAAGAGCCTCAGAGAGCTGGAAGCTGAGCCTGAAATGAG | 2607 |
| Db | 1138 | CTGGCTCTAACTCTCAGTGTGAACCAAGAGCCTCAGAGAGCTGGAAGCTGAAATGAG | 1197 |
| QY | 2608 | CTGGGGGACCTCGGGTGTGCTGTGTGTGAGGGCTCAGGCATCCCACTGTCGAAGCTC | 2667 |
| Db | 1198 | CTGGGGGACCTCGGGTGTGCTGTGTGTGAGGGCTCAGGCATCCCACTGTCGAAGCTC | 1257 |
| QY | 2668 | CAGACCTCGGGTTGGCATTCTCGGCTCGGCTCTGCGGCTGTGAGGCTCTTTCTGTG | 2727 |
| Db | 1258 | CAGACCTCGGGTTGGCATTCTGCGGCTGGGCTCTGCGGCTGTGAGGCTCTTTCTGTG | 1317 |
| QY | 2728 | GTGCTCAAGCCCAACCAACAACCTCCGGAGCTGGAATGTGATTTCAACGACCTGGGAGAC | 2787 |
| Db | 1318 | GTGCTCAAGCCCAACCAACAACCTCCGGAGCTGGAATGTGATTTCAACGACCTGGGAGAC | 1377 |
| QY | 2788 | TGGGGCTGTGGTTGCTGGCTGAGGGCTGCAACATCCCGCTTGCAGACTCGAGAACTG | 2847 |
| Db | 1378 | TGGGGCTGTGGTTGCTGGCTGAGGGCTGCAACATCCCGCTTGCAGACTCGAGAACTG | 1437 |
| QY | 2848 | TGGCTGGATAGCTGTGGCCTCACAGCCAAGGCTTGTGAGAACTTTTACTTTCACCCCTGGG | 2907 |
| Db | 1438 | TGGCTGGATAGCTGTGGCCTCACAGCCAAGGCTTGTGAGAACTTTTACTTTCACCCCTGGG | 1497 |
| QY | 2908 | ATCAACAGACCTTGACCGACCTTTACTTGACCAACAACGCCCTTAGGGGACACAGGTGTC | 2967 |
| Db | 1498 | ATCAACAGACCTTGACCGACCTTTACTTGACCAACAACGCCCTTAGGGGACACAGGTGTC | 1557 |
| QY | 2968 | CGACTGTTTGCAGGGCTGAGGCATCTTGGCTGCMAACTCCGAGTCCTCTGGTTATTT | 3027 |
| Db | 1558 | CGACTGTTTGCAGGGCTGAGGCATCTTGGCTGCMAACTCCGAGTCCTCTGGTTATTT | 1617 |
| QY | 3028 | GGGATGGAAGCTGAAATAAAATGACCCACAGTAGGTTGGCAGCGCTTCAGATAAACAAACCT | 3087 |
| Db | 1618 | GGGATGGAAGCTGAAATAAAATGACCCACAGTAGGTTGGCAGCGCTTCAGATAAACAAACCT | 1677 |
| QY | 3088 | TATTTGGAACATTGGCTGCTGA | 3108 |
| Db | 1678 | TATTTGGAACATTGGCTGCTGA | 1698 |

RESULT 6
ACD03624

ACD03624
ID ACD03624 standard; cDNA; 1683 BP.

AC ACD03624;

XXXX

DT 01-AUG-2003 (first entry)

XX DE Novel human GPCR related protein NOV2b cDNA.

XX
KW
Human; G-protein coupled receptor related protein; GPCR related protein;

NOV; cytostatic; cardiant; antiarteriosclerotic; antidiabetic; immunomodulator; anti-HIV; anorectic; antiasthmatic; haemostatic; antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine; NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer; diabetes; immune disorder; AIDS; obesity; asthma; haematopoietic disorder; Parkinson's disease; Alzheimer's disease; infection; multiple sclerosis; cancer-associated cachexia; wasting disorder; chronic disease; neurogenesis; cell differentiation; cell proliferation; haematopoiesis; wound healing; angiogenesis; chromosome mapping; tissue typing; preventive medicine; pharmacogenomic; gene; ss.

Homo sapiens.

WO200299116-A2.

12-DEC-2002.

04-JUN-2002; 2002WO-US017428.

04-JUN-2001; 2001US-0295607P.

04-JUN-2001; 2001US=0293661P;
06-JUN-2001; 2001US=0296404P;

06--JUN-2001; 2001US-0296418P.
14--JUN-2001: 2001US-0298285P

15-JUN-2001; 2001US-0298556P.
21-JUN-2001; 2001US-0388949P

26-JUN-2001; 2001US-0300883P.

13-AUG-2001; 2001US-0311972P.

29-AUG-2001; 2001US-0315660P.

17-SEP-2001; 2001US-0322706P.

28-FEB-2002; 2002US-0361189P.

12-MAR-2002; 2002US-0363676P.

0

1. THEORY

Anderson DW, Baumgartner UC, Gangolli EA, Gerlach VL, Go

MacDougall JK, Malyan KA, OM, Pena CE, Rastelli L, Shimk

VOSS, ET AL., ZEPHYRUS

WPI; 2003-140627/13.
P-PSDB: ABII99120

New NOX polymides and nitr

treating NOVX-associated diso-

pharmacogenomics.

Claim 20; Page 99-100

The invention describes an isolated polypeptide (I) comprising any of 27 118-961 residue amino acid sequences, given in the specification, a

or a sequence hav

for the first time, making it the first syndrome associated with a human disease selected from a pathology associated with the polypeptide. The NOVX polypeptides, polynucleotides associated with the polypeptide. The NOVX polypeptides, polynucleotides

and antibodies are useful in creating disorders, e.g. cardiomyopathy, atherosclerosis, etc.

diseases, AIDS, obesity, ascuma, haematopoietic disorders, Parkinson's disease, Alzheimer's disease, infections, multiple sclerosis, cancer-associated cachexia, and other wasting disorders associated with chronic diseases. The nucleic acids and polypeptides may also be used as targets for the identification of small molecules that modulate or inhibit e.g. neurogenesis, cell differentiation, cell proliferation, haematopoiesis,

CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods. The nucleic acids are further used as
CC hybridisation probes, in chromosome mapping, tissue typing, preventive
CC medicine, and pharmacogenomics. The polypeptides are also useful as
CC vaccines. This sequence encodes a novel human G-protein coupled receptor
CC related protein NOV
XX

Sequence 1683 BP; 343 A; 483 C; 508 G; 349 T; 0 U; 0 Other;

Query Match 39.0%; Score 1212; DB 7; Length 1683;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1212; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|------|---|------|
| QY | 1104 | GCCTCTCTTACCATGCTCTTCTGCTCCCTCGTGTGCTGGTGTGCTACTGCTCTCA | 1163 |
| DB | 5 | GCCTCTCTTACCATGCTCTTCTGCTCCCTCGTGTGCTGGTGTGCTACTGCTCTCA | 64 |
| QY | 1164 | GCAGCAGCTGGAGGGTGGGGGCTGTTTGAGACAGACGTCCAGGACCACTGTCAGTGT | 1223 |
| DB | 65 | GCAGCAGCTGGAGGGTGGGGGCTGTTTGAGACAGACGTCCAGGACCACTGTCAGTGT | 124 |
| QY | 1224 | CATGCTCTACCTGCTGAGTCTGATGCAACCAAGCCGGGGGCCCGCCCTCCAGCCCC | 1283 |
| DB | 125 | CATGCTCTACCTGCTGAGTCTGATGCAACCAAGCCGGGGGCCCGCCCTCCAGCCCC | 184 |
| QY | 1284 | ACCAACACAGAGAGGGTGTGCTCTCGCGCAGATGGGTCTCGAATCAGAAATCCT | 1343 |
| DB | 185 | ACCAACACAGAGAGGGTGTGCTCTCGCGCAGATGGGTCTCGAATCAGAAATCCT | 244 |
| QY | 1344 | ATTTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGAAGACGTCTCTGCTTCT | 1403 |
| DB | 245 | ATTTGAGGAGCAGGACCTCCGGAAGCAGCGCTAGACGGGAAGACGTCTCTGCTTCT | 304 |
| QY | 1404 | CAACATGAACATCTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCACTT | 1463 |
| DB | 305 | CAACATGAACATCTCCAGAGGACATCACTGTGAGAGGTACTACAGCTTCATCCACTT | 364 |
| QY | 1464 | GAGTTTCAGAAATCTTTGAGCTATCTATATCTGACGAGGGGGAGGGGGG | 1523 |
| DB | 365 | GAGTTTCAGAAATCTTTGAGCTATCTATATCTGACGAGGGGGAGGGGGG | 424 |
| QY | 1524 | AGGCCACAGACGAGGAGCTGTGTACCGAGTACGCGTCTTCTGAAAGAGGCTT | 1583 |
| DB | 425 | AGGCCACAGACGAGGAGCTGTGTACCGAGTACGCGTCTTCTGAAAGAGGCTT | 484 |
| QY | 1584 | CCTGGCACTACAGCCGCTCTGTTGGACTCTGAAAGGAGGAGCAGAGCCACT | 1643 |
| DB | 485 | CCTGGCACTACAGCCGCTCTGTTGGACTCTGAAAGGAGGAGCAGAGCCACT | 544 |
| QY | 1644 | GGAGAAGAGTCTCTGCGAAGTCTCGCGCACATCAAGATGGACCTGTTGAGTGGAT | 1703 |
| DB | 545 | GGAGAAGAGTCTCTGCGAAGTCTCGCGCACATCAAGATGGACCTGTTGAGTGGAT | 604 |
| QY | 1704 | CCAAAGCAAAAGCTCAGAGCGAGGCTCCACCTCTGACGAGGCGCTCTTGGAGTCTT | 1763 |
| DB | 605 | CCAAAGCAAAAGCTCAGAGCGAGGCTCCACCTCTGACGAGGCGCTCTTGGAGTCTT | 664 |
| QY | 1764 | CTGCTTGTACAGATCCAGAGGAGGAGTGTATCCAGAGGCGCTTCCAGGT | 1823 |
| DB | 665 | CTGCTTGTACAGATCCAGAGGAGGAGTGTATCCAGAGGCGCTTCCAGGT | 724 |
| QY | 1824 | GATCGTGTGTCAGCAATTCCTCCAGATGGAGCAGATGCTCTCTGTTCTGCTGAA | 1883 |
| DB | 725 | GATCGTGTGTCAGCAATTCCTCCAGATGGAGCAGATGCTCTCTGTTCTGCTGAA | 784 |
| QY | 1884 | GGCTGCGAGGAGCGCCAGGCTGCTGACTTGTATGGCGCCACTACAGCGGAGCGGGA | 1943 |
| DB | 785 | GGCTGCGAGGAGCGCCAGGCTGCTGACTTGTATGGCGCCACTACAGCGGAGCGGGA | 844 |
| QY | 1944 | AGACCGCGAGGCTGCTCCAGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACAGAG | 2003 |
| DB | 845 | AGACCGCGAGGCTGCTCCAGAGGAGCGCACAGCTGTTGGTGCAGCTCAGACAGAG | 904 |

RESULT 7

| | | | |
|----------|--|---------------|--------------------------|
| ACD03623 | ID | ACD03623 | standard; cDNA; 1800 BP. |
| XX | AC | ACD03623; | |
| XX | AC | ACD03623; | |
| DT | 01-AUG-2003 | (first entry) | |
| XX | Novel human GPCR related protein NOV2a cDNA. | | |
| DE | Human; G-protein coupled receptor related protein; GPCR related protein; | | |
| XX | NOV; cytotostatic; cardiant; antiarteriosclerotic; antidiabetic; | | |
| KW | immunomodulator; anti-HIV; anorectic; antiaesthetic; haemostatic; | | |
| KW | antiparkinsonian; neuroprotective; nootropic; gene therapy; vaccine; | | |
| KW | NOVX-associated disorder; cardiomyopathy; atherosclerosis; cancer; | | |
| KW | diabetes; immune disorder; AIDS; obesity; asthma; | | |
| KW | haematopoietic disorder; Parkinson's disease; Alzheimer's disease; | | |
| KW | infection; multiple sclerosis; cancer-associated cachexia; | | |
| KW | wasting disorder; chronic disease; neurogenesis; cell differentiation; | | |
| KW | cell proliferation; haematopoiesis; wound healing; angiogenesis; | | |
| KW | chromosome mapping; tissue typing; preventive medicine; pharmacogenomic; | | |
| XX | gene; ss. | | |
| OS | Homo sapiens. | | |
| XX | WO200299116-A2. | | |
| PN | 12-DEC-2002. | | |
| XX | 04-JUN-2002; 2002WO-US017428. | | |
| PD | 04-JUN-2001; 2001US-0295607P. | | |
| PF | 04-JUN-2001; 2001US-0295661P. | | |
| XX | 06-JUN-2001; 2001US-0296404P. | | |
| PR | 06-JUN-2001; 2001US-0296418P. | | |
| PR | 14-JUN-2001; 2001US-0298285P. | | |
| PR | 15-JUN-2001; 2001US-0298566P. | | |
| PR | 21-JUN-2001; 2001US-0299949P. | | |
| PR | 26-JUN-2001; 2001US-0300883P. | | |
| PR | 28-JUN-2001; 2001US-0301505P. | | |
| PR | 13-AUG-2001; 2001US-0311972P. | | |
| PR | 27-AUG-2001; 2001US-0315071P. | | |
| PR | 29-AUG-2001; 2001US-0315660P. | | |
| PR | 14-SEP-2001; 2001US-0322293P. | | |
| PR | 17-SEP-2001; 2001US-0322706P. | | |
| PR | 14-DEC-2001; 2001US-0341186P. | | |
| PR | 28-FEB-2002; 2002US-0361189P. | | |

RESULT 8

ABL59333
ID ABL59333 standard; cDNA; 4931 BP.

XX AC ABL59333;
XX DT 07-OCT-2002 (first entry)

Nucleotide sequence of human leucine-rich repeat protein HLRRBM1.

XX Human; leucine-rich repeat; HLRRBM1; proliferative disorder;
XX immune condition; apoptosis; signal transduction; autoimmune disease;
XX haematopoietic cell disease; graft-versus-host disease; allergy; asthma;
XX cardiovascular disorder; neurological disease; pheromone;
XX pulmonary disease; chronic obstructive pulmonary disease;
XX allergic rhinitis; bronchial hyperresponsiveness; reproductive disease;
XX haematopoietic disease; platelet disorder; Bernard-Soulier syndrome;
XX inflammatory disorder; systemic lupus erythematosus;
XX cardiovascular disease; cancer; gene; ss.

XX OS Homo sapiens.

XX Key Location/Qualifiers
XX CDS 1320..2666
XX /*tag= a
XX /*product= "HLRRBM1"

XX WO200252011-A2.

XX 04-JUL-2002.

XX 20-DEC-2001; 2001WO-US049740.

XX 22-DEC-2000; 2000US-0257773P.

XX (BRIM) BRISTOL-MYERS SQUIBB CO.

XX Feder J, Ramanathan C, Mintier G;

XX WPI; 2002-566676/60.

XX P-PSDB; ABB77910.

XX New HLRRBM1 nucleic acids for preventing, treating or ameliorating e.g.
XX proliferative disorders, immune conditions, a disorder related to
XX aberrant apoptosis modulation or developmental disorders.

XX Claim 1; Fig 1A-E; 371pp; English.

XX The present sequence encodes a human leucine-rich repeat containing
XX protein, designated HLRRBM1. HLRRBM1 polypeptides and polynucleotides are
XX useful for preventing, treating or ameliorating a medical condition such
XX as a proliferative disorder, immune condition, or a disorder related to
XX aberrant apoptosis modulation, either directly or indirectly, and in
XX modulating signal transduction activity in various cells, tissue and
XX organisms. They are also useful for treating, preventing, or diagnosing
XX diseases of haematopoietic cells, autoimmune disease, graft-versus-host
XX disease, allergic conditions (e.g. asthma), cardiovascular disorders, and
XX neurological diseases, and for increasing the organisms' ability to
XX synthesize and/or release pheromones. The polypeptide may also be used in
XX treating, preventing or ameliorating pulmonary disease (e.g. chronic
XX obstructive pulmonary disease, allergic rhinitis, or bronchial
XX hyperresponsiveness), reproductive disease, haematopoietic disease,
XX platelet disorders (e.g. Bernard-Soulier syndrome), non-infectious
XX disorders (e.g. innate immunity to bacterial pathogens, or adaptive
XX immune response), immune and inflammatory disorders (e.g. systemic lupus
XX erythematosus), cardiovascular diseases and cancers. HLRRBM1 nucleic
XX acids may further be used in chromosome identification or mapping, as a
XX chromosome marker, as molecular weight markers, as diagnostic probes, in
XX gene therapy, in raising anti-DNA antibodies, or as antigens for
XX eliciting immune responses

SQ Sequence 4931 BP; 1182 A; 1357 C; 1331 G; 1061 T; 0 U; 0 Other;

Query Match 24.7%; Score 768; DB 6; Length 4931;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | | | | | | | |
|----|------|-------------|----------|----------|----------|----------|----------|----------|------|
| QY | 676 | AGGAGATGAAC | CAGAGTGC | CCGGAATG | CAGATG | CAAGACCT | CATCTTC | CAGTGTGG | 735 |
| DB | 1314 | AGGAGATGAAC | CAGAGTGC | CCGGAATG | CAGATG | CAAGACCT | CATCTTC | CAGTGTGG | 1373 |
| QY | 736 | CTTGAGCCG | CGCGCTCT | CCAGAGCT | CATCGAGT | TCCGAGCG | CCCTCTT | CATC | 795 |
| DB | 1374 | CTTGAGCCG | CGCGCTCT | CCAGAGCT | CATCGAGT | TCCGAGCG | CCCTCTT | CATC | 1433 |
| QY | 796 | ATCGACGG | CTTCGATG | AGCTCA | AGCCCTT | TTCCACG | ATCCTC | AGGACCTT | 855 |
| DB | 1434 | ATCGACGG | CTTCGATG | AGCTCA | AGCCCTT | TTCCACG | ATCCTC | AGGACCTT | 1493 |
| QY | 856 | TGCTGGG | AGGAGAA | ACGGCC | ACGAGT | GTCTTT | AACAGCTT | AAATTC | 915 |
| DB | 1494 | TGCTGGG | AGGAGAA | ACGGCC | ACGAGT | GTCTTT | AACAGCTT | AAATTC | 1553 |
| QY | 916 | CTCCCTG | AGCTATCT | TTTGCT | CATC | ACCAC | AGCGCC | CTTTGG | 975 |
| DB | 1554 | CTCCCTG | AGCTATCT | TTTGCT | CATC | ACCAC | AGCGCC | CTTTGG | 1613 |
| QY | 976 | CTGCTG | GAGAC | CCCCAG | GCGATG | TGGAG | ATCTCTG | AGGCA | 1035 |
| DB | 1614 | CTGCTG | GAGAC | CCCCAG | GCGATG | TGGAG | ATCTCTG | AGGCA | 1673 |
| QY | 1036 | TACTTCT | ACAAGT | ATTTCC | ACAATG | CGAG | CGGCGC | CAAGTCTT | 1095 |
| DB | 1674 | TACTTCT | ACAAGT | ATTTCC | ACAATG | CGAG | CGGCGC | CAAGTCTT | 1733 |
| QY | 1096 | GACAA | CGAGC | CTCTCT | TCCATG | CTTCC | CTCCCTG | GTGTG | 1155 |
| DB | 1734 | GACAA | CGAGC | CTCTCT | TCCATG | CTTCC | CTCCCTG | GTGTG | 1793 |
| QY | 1156 | TGCTTC | CAGAC | GAGCTG | GAGCG | GTGGG | GGGTGT | TGAGAC | 1215 |
| DB | 1794 | TGCTTC | CAGAC | GAGCTG | GAGCG | GTGGG | GGGTGT | TGAGAC | 1853 |
| QY | 1216 | CGAGT | GATAC | TGCTCT | ACCTCT | GATG | CTGATG | CAAC | 1275 |
| DB | 1854 | CGAGT | GATAC | TGCTCT | ACCTCT | GATG | CTGATG | CAAC | 1913 |
| QY | 1276 | CAGCCCC | CAAC | CAACAG | AGGGT | TGCTC | CTTGG | CGGAG | 1335 |
| DB | 1914 | CAGCCCC | CAAC | CAACAG | AGGGT | TGCTC | CTTGG | CGGAG | 1973 |
| QY | 1336 | AAATCT | CTATT | TGAG | GACG | AGAC | CTCCG | GAAG | 1395 |
| DB | 1974 | AAATCT | CTATT | TGAG | GACG | AGAC | CTCCG | GAAG | 2033 |
| QY | 1396 | GCCTTC | CTCA | ATGA | CAATCT | TC | CAGA | AGGAC | 1443 |
| DB | 2034 | GCCTTC | CTCA | ATGA | CAATCT | TC | CAGA | AGGAC | 2081 |

RESULT 9

ADC32201

ID ADC32201 standard; cDNA; 591 BP.

XX AC ADC32201;

XX DT 18-DEC-2003 (first entry)

XX Human novel cDNA contig sequence, SEQ ID NO:2283.

XX Human; diagnostic; drug screening; forensics; gene mapping;

XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;

XX neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

XX ulcers; osteoporosis; autoimmune disease; cancer;

PT mapping, in the recombinant production of protein, or in generating
 PT antisense DNA or RNA.

PS Claim 1; SEQ ID NO 23437; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
 CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
 CC determined by the technique of SH (sequencing by hybridisation). Also
 CC included is a purified polypeptide comprising a sequence corresponding to
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences
 CC are useful in diagnostics as expressed sequence tags (EST) for
 CC identifying expressed genes or for physical mapping of the human genome,
 CC in forensics, in assessing biodiversity, or in identifying mutations
 CC responsible for genetic disorders and other traits. The nucleotide
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,
 CC for chromosome and gene mapping, in the recombinant production of
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide
 CC is useful for generating antibodies specific for it. The present sequence
 CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
 CC for this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from USPTO at
 CC seqdata.uspto.gov/sequence.html?docID=20030073623

XX SQ Sequence 487 BP; 97 A; 124 C; 155 G; 105 T; 0 U; 6 Other;

Query Match 12.4%; Score 385; DB 8; Length 487;
 Best Local Similarity 99.8%; Pred. No. 5.6e-165;
 Matches 435; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2212 GCAGCTCTCCTAGGCAATGAAGATTGACAGGATGATCTCAGTGGCAACGGGCTTGA 2271
 DB 31 GCAGCTCTCCTAGGCAATGAAGATTGACAGGATGATCTCAGTGGCAACGGGCTTGA 90
 QY 2272 TTCCAGGCATGATGCTCTTTGGAGGGCTGGGGCTCCAGTCCAGGCTGCAGATG 2331
 DB 91 TTCCAGGCATGATGCTCTTTGGAGGGCTGGGGCTCCAGTCCAGGCTGCAGATG 150
 QY 2332 ATTGAGTTGAGGAGTGTGAGTGGAGTCCGGGGCTCTCAGGAGATGGCTTCTGTGCTC 2391
 DB 151 ATTGAGTTGAGGAGTGTGAGTGGAGTCCGGGGCTCTCAGGAGATGGCTTCTGTGCTC 210
 QY 2392 GGCACCAACCCACATCTGTTGAGTTGAGCTGACAGGAAATGCACTGGAGGATTTGGGC 2451
 DB 211 GGCACCAACCCACATCTGTTGAGTTGAGCTGACAGGAAATGCACTGGAGGATTTGGGC 270
 QY 2452 CTGAGTTACTATGTCAGGAGTACGAGCCAGTCTGAGACTACGAGCTTTGGGCTG 2511
 DB 271 CTGAGTTACTATGTCAGGAGTACGAGCCAGTCTGAGACTACGAGCTTTGGGCTG 330
 QY 2512 AAGATCTCCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2571
 DB 331 AAGATCTCCGCTCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390
 QY 2572 CAGAGCTGAGAGAGCTGGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 2631
 DB 391 CAGAGCTGAGAGAGCTGGACCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 450
 QY 2632 CTGTGTGAGGGCTTCA 2647
 DB 451 CTGTGTGAGGGCTTCA 466

RESULT 11

AAC76566

ID AAC76566 standard; cDNA; 479 BP.

XX AAC76566;

AC AAC76566;

DT 08-FEB-2001 (first entry)

DE Human ORFX ORF2121 polynucleotide sequence SEQ ID NO:4241.

XX Human; open reading frame; ORFX; detection; cytotstatic; hepatotropic;

KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;
 KW anticonvulsant; osteoparathic; antiarthritic; immunosuppressant; cardiac;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease; AIDS;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

OS Homo sapiens.

XX WO2000058473-A2.

XX 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

XX 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42357.

XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 3436; 5507pp; English.

CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antipsoriatic; antiparkinsonian; nootropic; neuroprotective; osteopathic;
 CC anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 CC cardiac; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 CC dermatological; immunosuppressive; antiinflammatory; antibacterial;
 CC antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
 CC sequences can be used for determining the presence of or predisposition
 CC to, or preventing or treating pathological conditions associated with an
 CC ORFX-associated disorder. The nucleic acids can be used to express ORFX
 CC proteins in gene therapy vectors. The proteins and nucleic acids may be
 CC used to treat cancers, proliferative disorders, neurodegenerative
 CC disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 CC diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 CC storage, systemic lupus erythematosus, severe combined immunodeficiency
 CC (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 CC disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 CC cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 CC enhance coagulation; to inhibit thrombosis; and as a contraceptive
 XX Sequence 479 BP; 101 A; 139 C; 145 G; 93 T; 0 U; 1 Other;

Query Match 11.1%; Score 346; DB 3; Length 479;

Best Local Similarity 99.7%; Pred. No. 3.4e-147;

Matches 396; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1562 ACGCGTTTCTGAAGAGGAGCTCTCGGCACTCACCAGCGGCTTCTGTTGGACTCTGA 1621

DB 2 ACGCGTTTCTGAAGAGGAGCTCTCGGCACTCACCAGCGGCTTCTGTTGGACTCTGA 61

QY 1622 ACGAGGACACGAGGACCTGGAGAGAGTCTCTGCTGGAAGGTCTCGGCGCATCA 1681

CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIFO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 3306 BP; 718 A; 958 C; 939 G; 691 T; 0 U; 0 Other;

Query Match 9.3%; Score 289; DB 9; Length 3306;
 Best Local Similarity 100.0%; Pred. No. 3e-121;
 Matches 289; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGGAGGACCGCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Db 497 ATGCTACGAACCGCAGGAGGACCGCTCTGTGCGCTGTCCCTACTTGGAGAACTC 438
 Qy 61 GAGCGTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGAGCGGACAGAGCTGGAGAA 120
 Db 437 GAGCGTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGAGCGGACAGAGCTGGAGAA 378
 Qy 121 GCGAAGATCCCTGGGAGGATGAGAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Db 377 GCGAAGATCCCTGGGAGGATGAGAGGCGGCTCCCTGGAATGGCCAGCTGCTC 318
 Qy 181 ATCACCACCTTGGGCGCAGAGAGGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
 Db 317 ATCACCACCTTGGGCGCAGAGAGGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 258
 Qy 241 AACAGGAGGACCTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 289
 Db 257 AACAGGAGGACCTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 209

RESULT 14

AAK50424

ID AAK50424 standard; DNA; 282 BP.

XX AC AAK50424;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 24981.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

OS

XX WO200157276-A2.

PN

PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US0000668.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-489900/53.

DR Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human bone marrow.

XX Example 4; SEQ ID NO 24981; 658pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention

XX Sequence 282 BP; 71 A; 71 C; 96 G; 44 T; 0 U; 0 Other;

Query Match 9.1%; Score 282; DB 4; Length 282;
 Best Local Similarity 100.0%; Pred. No. 5.4e-118;
 Matches 282; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGCTACGAACCGCAGGAGGAGCGCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Db 1 ATGCTACGAACCGCAGGAGGAGCGCTCTGTGCGCTGTCCCTACTTGGAGAACTC 60
 Qy 61 GAGCGTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGAGCGGACAGAGCTGGAGAA 120
 Db 61 GAGCGTGTGGAACTGAAGAAGTTCAAGTTATACCTGGGAGCGGACAGAGCTGGAGAA 120
 Qy 121 GCGAAGATCCCTGGGAGGATGAGAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Db 121 GCGAAGATCCCTGGGAGGATGAGAGGCGGCTCCCTGGAATGGCCAGCTGCTC 180
 Qy 181 ATCACCACCTTGGGCGCAGAGAGGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
 Db 181 ATCACCACCTTGGGCGCAGAGAGGCGCTGGAGTTGGCTCTCAGACCTTTGAGCGGATA 240
 Qy 241 AACAGGAGGACCTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 282
 Db 241 AACAGGAGGACCTTGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 282

RESULT 15

AAK50424

ID AAK50424 standard; cDNA; 506 BP.

XX AC AAK50424;

XX 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #4561.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX P-PSDB; ABG04570.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 1; SEQ ID NO 4561; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS04197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 506 BP; 108 A; 149 C; 148 G; 101 T; 0 U; 0 Other;
Query Match 8.2%; Score 254; DB 5; Length 506;
Best Local Similarity 100.0%; Pred. No. 3.1e-105; Indels 0; Gaps 0;
Matches 254; Conservative 0; Mismatches 0;
Qy 1614 ACTCTGAACGAGGAGACCCAGGAGCCACCTGGAGAGAGTCTCTGTGGAAGGTCTCGCC 1673
Db 33 ACTCTGAACGAGGAGACCCAGGAGCCACCTGGAGAGAGTCTCTGTGGAAGGTCTCGCC 92
Qy 1674 GCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCAC 1733
Db 93 GCACATCAAGATGGACCTGTTCAGTGGATCCAAAGCAAGCTCAGAGCGGCTCCAC 152
Qy 1734 CTGTGACGAGGCTCTCTGTGAGTTCCTTCAGTCTGTTCAGATCCAGGAGGAGGTT 1793
Db 153 CTGTGACGAGGCTCTCTGTGAGTTCCTTCAGTCTGTTCAGATCCAGGAGGAGGTT 212
Qy 1794 TATCCAGCAGGCGCTGAGCCACTTCCAGGTCATCGTGTGATCGCAACATTCCTCCAAAGAT 1853
Db 213 TATCCAGCAGGCGCTGAGCCACTTCCAGGTCATCGTGTGATCGCAACATTCCTCCAAAGAT 272
Qy 1854 GGAGCACATGGTCT 1867
Db 273 GGAGCACATGGTCT 286

Search completed: August 8, 2004, 10:45:49
Job time : 1135 secs